FInest – **F**uture **In**ternet **e**nabled optimi**s**ation of **t**ransport and logistics networks

D5.1

Requirements Analysis and Selection of Technology Baseline for Collaboration Manager

Project Acronym	FInest	
Project Title	Future Internet enabled optimisation of transport and logistics networks	
Project Number	285598	
Work Package	#WP5 Business Collaboration Network	
Lead Beneficiary	SAP	
Editor(s)	René Fleischhauer	SAP
Contributors(s)	Cyril Alias	UDE
	Stephan Heyne	SAP
Reviewer	Fabiana Fournier	IBM
Reviewer	Øyvind Olsen	NCL
Dissemination Level	Public	
Contractual Delivery Date	30th September 2011	
Actual Delivery Date		
Version	V1.0	

The research leading to these results has received funding from the European Community's Seventh Framework Programme [FP7/2007-2013] under grant agreement no. 285598

The research leading to these results has received funding from the European Community's Seventh Framework Programme [FP7/2007-2013] under grant agreement no. 285598

Abstract

This document contains the first deliverable of Work Package 5. The work package is responsible for the Business Collaboration Module (BCM), which aims at the introduction of an infrastructure to securely manage the end-to-end networks of transport and logistics partners. It integrates information from different external sources as well as other modules of the FInest platform and makes this available for end-users of the system. In order to ensure the non-disclosure of confidential data, the BCM enables user and access management, which provides users with specific views on the data accordingly to their individual disclosure level.

This document describes the currently applied ICT systems for collaboration in the T&L domain. Based on this, an initial draft of the BCM's conceptual architecture is presented, which aims to solve previously determined drawbacks of current systems. This conceptual architecture is basis for a preliminary identification of technical and functional requirements and the selection of baseline technologies, which follows in the subsequent sections of the document. A final section aims at the alignment of the BCM to the FI PPP Core Platform and identifies potential Generic Enablers, which can be used to implement the envisioned features.

Consequently, this document addresses two of the tasks, defined in the Description of Work, for Work Package 5: T5.1 – Requirements Analysis and Selection of Technology Baseline – and partially T5.3 – Technological Alignment with the FI PPP Core Platform. In addition to this, it provides a description of the conceptual architecture of the BCM and partially addresses T5.2 – Conceptual Design and Technical Specification – with this.

Document History

Version	Date	Comments
V0.1	10-08-2011	First draft
V1.0	16-09-2011	Content completed; ready for internal review
V1.5	27-09-2011	Reviews integrated; ready for submission

Table of Contents

List of Tables

List of Figures

Acronyms

Acronym	Explanation
BCM	Business Collaboration Module
CO	Collaboration Object
C2K	Cargo 2000
ECM	E-Contracting Module
EDI	Electronic Data Interchange
EMF	Eclipse Modeling Framework
EPM	Event Processing Module
ERP	Enterprise Resource Planning
ETA	Estimated Time of Arrival
FI	Future Internet
GE	Generic Enabler
GPS	Global Positioning System
IATA	International Air Transport Association
JSON	Java Script Object Notation
OEM	Original Equipment Manufacturer
RDF	Resource Description Framework
RDF/S	RDF Schema
RFID	Radio Frequency Identification
T&L	Transport and Logistics
TAM	Technical Architecture Modeling
TPM	Transport Planning Module
UML	Unified Modeling Language
XML	eXtensible Markup Language

1. Introduction

Modern logistics processes are usually highly distributed, involving various stakeholders, with very different tasks and business models. Basically, such a process aims at the delivery of a good from a consignor to a consignee, but in fact this often results in the crossing of multiple borders, the change of continents, and the use of a variety of different transport vehicles (e.g., truck, train, vessel, or air plane). Thus, a logistics process can be considered as a distributed process with numerous steps, which all have to be executed successfully to finally deliver the goods to their destination. These steps are denoted as transportation legs (short: legs) and different stakeholders take over the responsibility for one or more of these. Moreover, it is possible that an involved party sub-contracts additional partners for its own leg. Apart from the stakeholders which are dealing with the physical transfer of the good, there are also legal parties (e.g., insurances) and governance authorities (e.g., border control or customs) are involved.

The different stakeholders depend on each other, whereas everybody needs detailed process information of the shipment and general information about the transferred goods. For instance, if a transport leg is delayed by an unforeseen event, the following party in the process needs to be informed in order to react accordingly. Another example is that a customs control can be tremendously accelerated if information about the shipped goods is available before a vessel reaches the port. However, the coordination between the different involved stakeholders is currently a highly manual process. Current status information between stakeholders is usually transmitted and received by humans via informal notifications (e.g., hard copied documents, telephone, fax, or email). Although most of the big logistics and transportation companies have proprietary ICT systems to facilitate the intra-organizational handling, these systems are mostly not able to breach the corporate borders. Hence, there is currently no ICT support available that enables a global view on the logistics process with customization support for the individual stakeholders and their different roles. This also can be described by a lack of an end-to-end view which means that all stakeholders, that currently are not handling the transport good, are not able to receive actual process information. This situation hampers an efficient planning and re-planning due to the fact that a lot of manual effort is required to omit the necessary information.

The *Business Collaboration Module* (short: BCM) introduces an infrastructure to securely manage end-to-end networks between transport and logistics partners. Its main task is the execution of a transport plan created by the *Transport Planning Module* (TPM) by the integration of information of other FInest modules – such as the, the *E-Contracting Module* (ECM) and the *Event Processing Module* (EPM) – and external legacy systems (e.g. ERP) as well as user input. During the execution of logistics processes the BCM which provides all involved stakeholders with the necessary information to conduct the process and reflects the actual process states by the integration of up-to-date events from the EPM. All information relevant to a specific logistics process is kept in an actual managed storage, with access-control and provides customized views on the data for each involved stakeholder. For this, the BCM uses so called *Collaboration Objects* (CO), whereby each encapsulates all information of a process's aspect. Through the connection of different COs, the whole logistics process is described. This approach provides the BCM with the necessary flexibility to adapt the used data model to changes that occur either at the runtime or at the planning phase of a logistics process (as will be explained in Section 4.1).

In comparison to existing Business-2-Business (B2B) used in the T&L domain, the BCM is built on top of them. It integrates existing systems B2B, like the SAP Business-to-Business Procurement solution or existing Supply Chain Management systems (e.g., from SAP, Oracle), and makes, among other things, the provided information available for external users. A major issue of current ICT solutions within the T&L domain is their heterogeneity. Different logistics companies use different, often incompatible, systems and this hampers an effective information exchange if they are participants in one logistics process. The BCM solves this problem by the provision of a global platform and the integration of different B2B systems. This allows a re-use of well-established infrastructure.

The remainder of this document is structured as follows: In the first chapter we consider the stakeholders, current systems, and methods for collaboration within the logistics and transport domain. For this, we give an overview about the different involved stakeholders and introduce currently used systems. In a subsequent chapter we present an initial conceptual architecture of the BCM, which is the basis for the identification of technical and functional requirements. With regard to the identified requirements, we conduct an initial technical State-of-the-Art analysis in the following section in order to determine suitable technologies or research approaches to implement the required features of the BCM. Additionally, we provide a first identification of potential Generic Enablers that reveals applicable technologies of the FI PPP Core Platform. In the last chapter, we provide the preliminary selection of a technology baseline for the BCM.

2. Current Technologies for Collaboration in the T&L Domain

In this chapter we investigate the current conduction of collaboration between the different stakeholders of logistics processes. With this we reveal potential improvements, which can be addressed by the BCM. In a first section we present a standard logistics process and explain the collaboration of the different involved stakeholders. This shows the complexity of modern logistics processes and gives a better understanding of the occurring problems. Subsequent to this, we present different types of currently applied ICT systems for collaboration and describe concrete systems as examples. In a last section, we denote the current downsides of the current ICT solutions and therewith, introduce the current issues which have to be addressed by the BCM.

2.1. Stakeholder Collaboration in a Standard Logistics Process

Transporting goods from seller to buyer usually involves several parties whose roles may be unclear to many external contemplators. In the following, we present an outline the basic functions of the most prominent actors in a transport chain. We omitted a definition of the used terminology, due to the fact that the provision of domain dictionary is done by Work Package 1 in Deliverable D1.2.

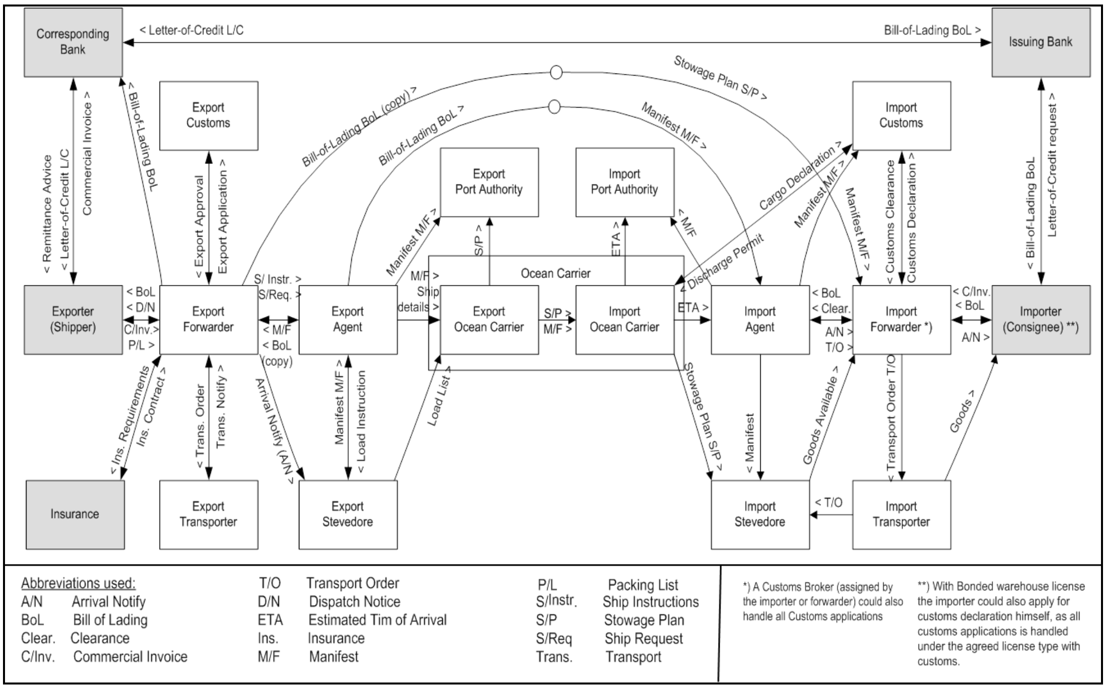

Figure 1 - IBM and KN Secure Trade Lanes Project

As can be seen in Figure 1 above, a multitude of parties are involved in an international transportation process. In terms of business collaboration, a large amount of documents and information have to be exchanged in order to execute the entire process smoothly. In order to show the complexity of the information flow accompanying the actual material flow, both the information received from and sent to the different players in a transport chain network are to be presented separately. In this way, each and every partner can be presented individually regarding his obligations in terms of information flow - which again highlight the plurality and complexity of interconnections.

This part intended to highlight the complex interconnections between the different players involved in business collaboration, which a large amount of documents and information have to be exchanged in. A complete list of all exchanged documents can be found in the appendix in Chapter 10.

Generally speaking, *Supply Chain Collaboration* basically follows the goal of pursuing self-interest and the own initiative of an individual member of a supply network and to support the goal by providing appropriate infrastructure, information and processes. In such a supply chain collaboration, decisions are made decentralized and voluntarily, and the partnerships do not follow a certain rule but make the entire supplier network of an automotive Original Equipment Manufacturer (OEM) available for collaboration [1].

However, since collaboration has to follow a regular network, the provision of a pertaining network infrastructure and of processes based thereupon is mandatory for the effective execution. A supply network management instance thus has to initiate and plan supply chain collaboration, then implement the latter by creating and providing the network base, and lastly extend and optimize the collaboration in case of successful usage or replace the ineffective collaboration entities [1].

Depending on the degree of collaboration as well as on the potential and the efforts envisioned, different levels of supply chain collaboration can be attained. Major factors behind the decision for a specific level of collaboration are the increased transaction costs and decreasing supply network costs. The optimal level of collaboration is where transaction costs are as high as supply network costs. As can be seen in the following, from mere information over different ways of planning to strategic orientation, supply chain collaboration can refer to various degree level [1].

With a relatively low degree of collaboration and low potential, transparency can be realized. Transparency can refer to the mere depiction of the network and the supply relationships therein but can also be raised to sophisticated information flow illustrations. The next level of supply chain collaboration is about alignment and coordination which again requires a structured approach as much as a certain level of synchronization. Collaborative Planning is an example for this level. The next higher level is the level of optimization over the entire supply network, in which planning and optimization of each individual partner within a supply network is coordinated transparently. The highest level of supply chain collaboration refers to strategic planning for which structures have to be planned, prepared, and provided collaboratively [1].

Regardless of the level of collaboration, each of the information exchange and communication tasks that need to be carried out by the different players take place in different systems and IT environments, if at all. The following section will deal with the current situation in logistics with regard to the use of ICT in a supply network for collaboration purposes [1].

2.2. Currently Applied ICT Systems for Collaboration (As-is Analysis)

Currently, the collaboration between different logistics partners is hardly realizable only by the use of a few ICT systems (or even just one). In order to accompany one and the same goods flow in current transportation processes the inclusion of a multitude of ICT systems along the supply chain is inevitable. Moreover, the **level of automation is highly heterogeneous** – both with regard to the parties involved and to the respective scope of use. For instance, there exist various systems that are envisioned for narrow purposes only. There exist specialist solutions for documents handling, procurement and purchasing, goods entrance and exit, warehouse management, tracking and tracing, management of receptacles like pallets, customs handling, financial management, insurance brokerage, transport orders and bookings, and fleet and transport management – just to mention a few. The sources of such systems are also highly heterogeneous, including data from both technical elements such as GPS, RFID and sensor networks, and organizational sources like HR time sheets, vehicle maintenance, and purchasing developments. The inclusion of such information is often limited to one or few destination systems though. That means they are used by one party only, not by several players involved in the actual process.

Despite the multitude of systems in use in a company, there still exists additional information that cannot be integrated into the systems and thus comes via conventional communication means, such as e-mail, telephone, facsimile or even mails and paper documents. All this leads to a considerable need for input, mostly manually or following personal approval by an employee.

Moreover, there exists a **large variety of solutions for each kind of systems** mentioned, each of them slightly differing from one another. Oftentimes, they are tailor-made for the needs and purposes of a particular company, business unit, subsidiary or even department. While large software packages oftentimes miss the core of the niche business models of many SMEs, self-made and tailored IT solutions are said to be able to contribute significantly to the efficiency in the respective process, albeit frequently not to other partners in the supply network. Adding to the heterogeneous picture and thus to the problem, variety also exists regarding both whose system to use. Depending on the intensity and quality of a business relationship or due to practical reasons, the decision in favour of a particular system is made. However, this adds to the problem of lacking interoperability between the ICT systems in a supply network.

For instance, several of the FInest domain partners indicated during the interview phase that they use own tailor-made (and often even self-made) solutions in order to facilitate tracking and tracing. Instead of automated data collection and update, the systems frequently base on manual entries and approval or on data input from EDI files, at best. The scope of the systems varies from pure notification (of departure from origin and arrival at destination and of goods loading and unloading, respectively) to continuous tracking of movements at each station. Yet, a continuous tracing of a logistics object has rarely been offered in such systems.

The **connection to a company's superior ERP system** is managed by means of interfaces or even manual transfer of data and information. Generally, the amount of work conducted manually or non-automatically is still on a noteworthy level - including the use of telephones, e-mails, letters and table sheets. Especially in the dealing with governmental and international bodies, such non-automated information transfers are rather the rule than the exception.

The FInest domain partners confirmed the situation depicted since most of their systems are not connected to their ERP system. It even happens as a rather usual case that the information gained in one system has to be re-entered in other systems again.

In many cases, the **use of unique identifying numbers** for a type of objects, e.g. for purchase orders or transport items, is not used throughout a company, let alone the supply network. That means, such reference numbers are used by the respective departments for internal purposes only whereas upstream and downstream parties in the supply network would possibly have to collect the same information once again. One of the domain partners in the FInest project indicated during the interview phase, the use of such unique numbers confines to the monitoring and the identification of document of purchasing orders.

Same applies to **communication systems** which mainly are not included in many of the systems mentioned before. Some systems do encompass the communication feature, but are limited to the interaction between two parties only. While some systems do allow (semi-)automated communication and information transfer, some do not offer the communication features required. Commonly, existing IT systems are 'misused' for communication purposes due to the mere lack of a proper communication system. For this reason, telephone and e-mail correspondence is carried out in order to ensure the information flow running in parallel to the material flow.

Apart from the material flow, the parties involved have to be interconnected in an informational manner. Basic **Information systems** provide the relevant companies with the necessary information for a particular purpose. As has been presented earlier in this document, the parties involved are of a highly varying nature which again is reflected in the IT systems used. For instance, many freight forwarder and carrier offer websites for tracking information. Likewise, for the different transportation modes, there exist certain information services such as CESAR (Co-operative European System for Advanced Information Redistribution) for combined traffic, ELWIS (Electronic Waterways Information System) for maritime transportation and Inforwarding.com for the air cargo business sector. These systems offer departure and arrival dates, tracking information, legal regulations, and traffic information, amongst others ([2], pp. 336-337).

Communication systems do not merely provide information for the transportation process in a passive manner, but actively support the flow of both goods and information. They can be differentiated between data collection and transmission systems, data exchange systems and data aggregation systems ([2], p. 337).

Data Entry and Transfer systems serve the electronic collection and transfer of information, e.g. on consignment and traffic. An automated processing of the information including the transmission of these to further systems is possible, but oftentimes not realized. To this category belong computer reservation systems for cargo consignments like GF-X (Global Freight Exchange) and CPS (Cargo Portal System) ([2], pp. 337-338).

Data Exchange Systems facilitate common data processing across the supply chain by the collection and integration of consignment information into an ICT system. In the transport and logistics domain, the standards IATA-CASS (IATA Cargo Accounts Settlement System) and HERMES (handling through European Railway Message Electronic System) have gained in importance in their respective goods transport sectors ([2], p. 338). Moreover, there exists so-called Cargo Community Services (CCS) which include features of data collection, data transmission, data exchange and general information management and which operate on a common platform by means data processing networks of many different actors of a transport or logistics chain. CCS are large integrating software packages consisting of many dedicated modules and can be classified as bound to either the mode of transport or the location ([2], pp. 338-339). Examples for CCS bound to the transport mode are TRAXON (Tracking and Tracing Online) for the air cargo industry and INTTRA, GT Nexus and CargoSmart for water-borne transportation. Examples for CCS bound to the location are DAKOSY (Hamburg, Germany), debit IT Services Benelux (Rotterdam, The Netherlands), SEAGHA (Antwerp, Belgium) and CNS (Southampton, UK) ([2], pp. 339-343).

One example for a Cargo Community Services is Cargo2000 (C2K), an airfreight management system launched by IATA. C2K defines quality standards for the supply chain, aims at improving the efficiency of the air cargo industry, improving customer Service and reducing costs to all participants by implementing a program of agreed business and automation standards, which are measurable and lead to quality competitive performance [3-5]. According to IATA, Cargo 2000 is a quality standard for the supply chain aiming at an improvement of the efficiency of the air cargo industry, of customer service and cost reduction to all participants by implementing a program of agreed business and automation standards, which are measurable and lead to quality competitive performance [3-5]. IATA claims as main objectives the attainment of the transparency and visibility of the actual goods flow, of the improvement of internal processes to enhance service delivery, of the recognition of adherence to certain quality standards, to mention only some major goals [3-5].

2.3. Known Issues of Currently Applied ICT Systems for Collaboration

Generally speaking, collaboration exists for different single or few transactions only. These solutions are mostly insular and non-adaptive solutions of the companies with a lot of restrictions, such as the limitation to certain process steps, the ability to integrate few parties only, and inflexibility to dynamic changes.
The major downsides of existing systems range from their inapplicability for complete supply chains and networks. Thus, there is no integration and communication between throughout the entire chain existing. The communication with the different stakeholder systems is to be harmonized in the future in order to facilitate a smooth information flow between the partners.

Oftentimes, manual entries are vital for the initiation of processes in the ICT systems. Especially, in those systems that actually are to support human decision-makers with their decision support information exhibited presently require manual input of the persons. In that regard, the quality and consistency of the data – currently a massive problem in the domain – have to be maintained by novel systems.

Thus, a real-time tracking is not possible since the relevant data is fed ex post into the respective systems. Likewise, a majority of systems is inappropriate for alert purposes. Moreover, delays based on lacking transparency regarding the tracking of the original documents are inevitable.

The FInest domain partners that have contributed to this report stated that oftentimes, communication between two partners takes place via conventional means mentioned earlier. If ICT systems are used at all, the communication is restricted to only two parties.

Although nowadays already several improvements exist, a large portion of the communication tasks is executed in a conventional manner, i.e. with paper, pencil and phone, as has been stated by the FInest domain partners. Interestingly, the identification number of purchasing orders, for example, is not used for the communication with other stakeholders of the process although the consistent use of such a unique identifier throughout the supply network highly appears to be useful. Another major communication approach is based on the use of EDI, which some of the domain partners use for message exchange with other partners of the supply network. The purpose of the use of EDI in this matter is the integration of this information in the very own system of the company which was impossible in the original data format. The circumvention of such issues can work both process-wise like agreement on standards and based on adjustments in IT systems following manual entries.

Finally, it has to be highlighted that a company involved in a transportation process has to deal with both the multitude of own systems and the systems of suppliers and customers. This is the major problem addressed by the BCM and hence, the issues of interoperability and the facilitation of the latter play a major role which again require a deep understanding of the respective data structures and IT system architectures. Accordingly, the FInest domain partners indicated that they have to deal with a lot of ICT systems along the supply network, mostly with a large number of manual entries and updates. For some of the domain partners, it even is a criterion for the selection of suppliers.

3. Conceptual Technical Architecture of the Business Collaboration Module

In this chapter we present the initial conceptual architecture of the BCM. This is the basis for the subsequent technical requirements analysis, the technical State-of-the-Art analysis and therewith, the selection of technology baseline. We use the previously described drawbacks of currently applied ICT systems, which can basically be summarized by the lack of a global view of all the information connected with logistics processes, and present our overall approach in the first section. In a subsequent section we present the BCM's initial conceptual architecture and describe the different components.

3.1. Approach and Overview

Within the FInest platform, the BCM executes previously created transport plans, provides all involved stakeholders of T&L processes with all necessary information and represent the current state of the process by an integration of event data. To implement this vision an appropriate approach to model data is a key requirement. In this section we describe the general modeling approach of the BCM and provides an illustrative example.

As mentioned in the introduction (cf. Chapter 1), a logistics process entail a wealth of different information. Each stakeholder requires a certain aspect of this information set in order to conduct their steps in the overall process. The role of a stakeholder (e.g. carrier, consignee or consignor) determines which information is needed and, more importantly, which security level a stakeholder has to access information.

In general, two different types of information within logistics processes can be distinguished. The first is all static data that is valid over the whole execution phase. Examples for this are the dimension and amount of transport goods, origin, destination or the ETA of the transport, contract information or data from existing ERP systems. The second type of information describes the current progress of the transport and is denoted as process information. While the first kind of information can be considered as static data, originating from (relational) data bases for instance, the second kind reflects the current state in a business process, which is changed during the execution of the process. Both kinds have to be managed by the BCM in an effective and efficient manner in order to provide end-users sufficient level of performance in real-world application scenarios which contain large amounts of data.

The overall data modeling approach of the BCM encapsulates different aspects of logistics processes in different so-called *Collaboration Objects* (CO). Each object combines the related static and process information in a single unit and the complete process is described by an interrelation of these objects. In this way, one possible realization of a description of logistics processes could be a graph, where COs are the nodes and connections between these objects implement graph edges. The scientific foundation for the idea is provided by the entity-centric modeling approach. This approach is developed by IBM and was initially introduced in [6]. A detailed examination of entity-centric modeling is provided in Section 5.2, which also assesses the applicability of this approach for the BCM.

At this stage it is important to understand that the concept of COs addresses the internal representation of data rather than beginning an exchange format. Hence, COs are primarily used to organize static and process data of logistics process in an effective and natural manner. Nowadays, a wealth of standardized data format for information exchange exists in the T&L domain. Probably, two of the most prominent representatives are EDIFACT[1] and the Cargo2000 [3] (for further examples please refer Figure 2 below or use Deliverable D2.2) . Although, the BCM is designed to support these well-established standards, the concept CO-based modeling does not address such standards directly. In order to support current communication standards the data representation within the COs will be converted to a specific format. The benefit of this approach is that multiple formats can be supported, depending on the preferences or capabilities of a requesting client. For this it is necessary that a CO contains all the vital information to generate the designated formats.

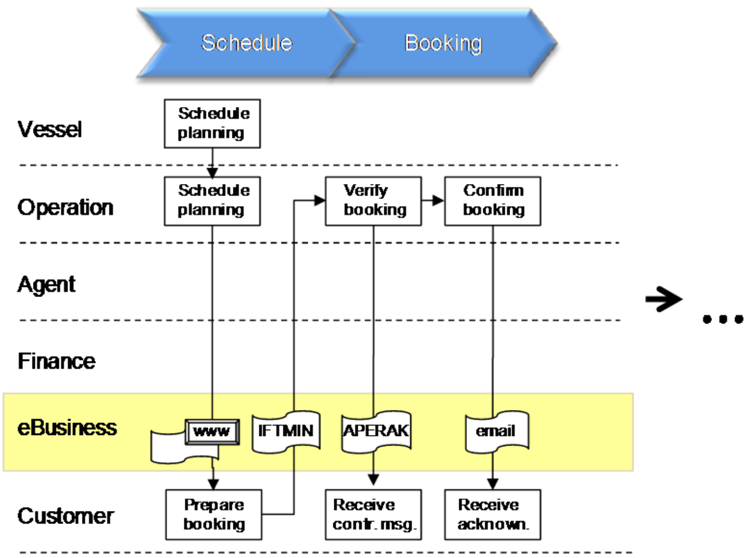

Figure 2 – Part of NCL e-Business “Schedule-to-Cash” flowchart (cf. Deliverable D2.2)

In order to explain the usage of COs in more detail, Figure 2 shows a part of a flowchart provided by NCL in Deliverable D2.2. It depicts the different steps of maritime container transports (implemented by NCL). Starting with scheduling plans for vessels, a customer selects a suitable date which fits his or her time constraints and a vessel which provides enough space to ship the designated transport good. The moment where the customer selects a certain option

[1] http://live.unece.org/trade/untdid/welcome.htm

and submits his or her request, a purchase order with a unique identifier (ID) is created. This purchase order has to be verified and confirmed in the subsequent process steps in order to implement the shipment (*Note:* the remaining process steps will be omitted here – cf. Deliverable 2.2 for further details). The creation of CO instances is in accordance to this process. If the user submits his or her booking request, an instance of the CO “PurchaseOrder” is created with its unique ID and time, origin, and destination as static data.. The process data encompass the possible states of the CO “PurchaseOrder” and would be initially set to “created”. The CO instance is now passed to the other involved stakeholders of the process, which accesses its (static) data and modifies its internal process status if needed. In the presented example, the operator would check the purchase data contained in the CO “PurchaseOrder” instance and set the internal process status from “created” to “verified”. Analogous to this the process status will be set to “confirmed” if the order is confirmed by the operator.

The BCM enables a global view of logistics processes, which means that every stakeholder can access all relevant information of the entire end-to-end process, for him or her, at a single point. To implement this, the BCM stores all data regarding a certain logistics process within a logically (physical location is hidden and therefore, not relevant) central point and makes it available for its users. Figure 3 illustrates this. Different stakeholders are connected via a Communication Layer with the BCM. Within the module exists a specialized storage component, the *Collaboration Object Storage*. As already mentioned, a logistics process is described by the means of *Collaboration Objects* composed in a graph based manner. The *Collaboration Object Storage* gathers many of these process descriptions and makes the contained data available for the connected stakeholders.

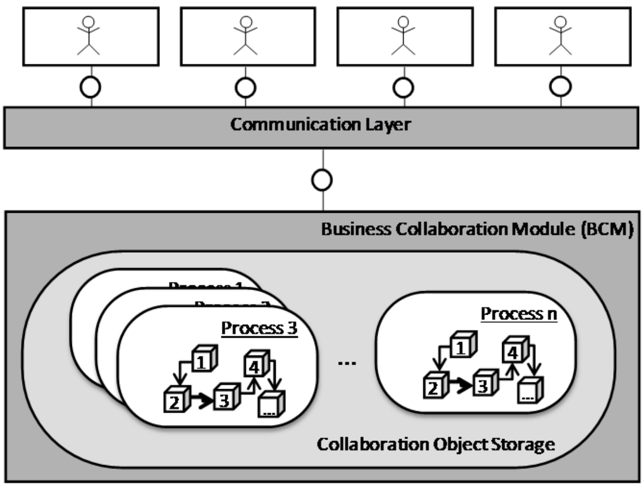

Figure 3 - Storage of various logistics process description in the BCM

3.2. Initial Conceptual Design

We present the initial conceptual design of the BCM in this section. Based on this, in subsequent sections we will identify technical requirements for the different internal components.

Figure 4 shows the conceptual design by the usage of the Technical Architecture Modeling (TAM) notation. The BCM itself is illustrated by the big gray rectangle which gathers the other components. Through different interfaces, depicted as small circles outside the rectangle, the BCM is able to interact with other modules of the FInest platform (for a precise description of the different modules please refer Deliverable D3.1). The shown interfaces are derived from the High-Level-Architecture introduced in Deliverable D3.1 and here not subject of further explanations. The following subsections describe the different internal components in detail.

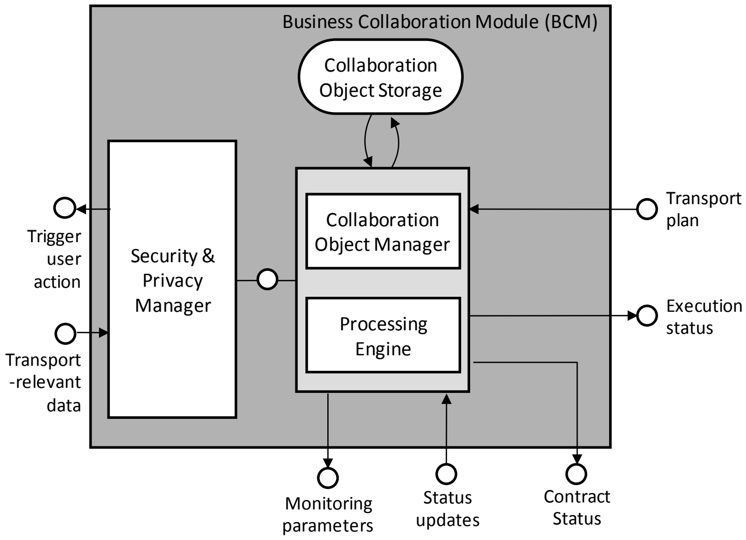

Figure 4 – Business Collaboration Manager (BCM) conceptual architecture

3.2.1. Security and Privacy Manager

Logistics processes entail a lot of confidential information, which have to be available for a certain set of stakeholders but not to others. For example, a hired carrier for a particular transport leg can sub-contract third party organizations to implement the physical transport. Although the original carrier does not reveal this fact to his own contract partner, the sub-contracted partners require information provided by their employers.

The *Security and Privacy Manager's* main task is to ensure the integrity of the stored data and the non-disclosure of confidential information. It ensures that certain information is available for the sub-contracted parties but not for the contract partner of the originally hiring carrier, for instance. It controls the information access in the BCM, user interfaces, and (legacy) third party systems. For this, it has to check whether information goes only to allowed users or systems.

3.2.2. Collaboration Object Manager

The BCM has to manage information originating different modules of the FInest platform as well as from the attached legacy systems from the backend layer. This information is composed of different data types, for example, the user can input coarse-grained information of a logistics process - like origin, destination or date - via a user interface, whereas detailed information about the goods are served by a backend ERP system. All kinds of this data have to be integrated into the BCM as well as stored, loaded and updated in order to provide the different end-users with the desired information.

The *Collaboration Object* Manager is the central data management unit in the BCM and is concerned with the instantiation of the data model artifacts based on all available information of a particular logistics process, the persistent storage of this data, and the retrieval of stored information. Thus, it controls the data of the actual logistics process and adapts it to the current status of the process or changed circumstances. By this, users of the BCM are enabled to retrieve up-to-date information and extract all required information.

3.2.3. Processing Engine

Most of the logistics processes consist of several stages, whereas each has to be implemented in order to execute the transport completely. The execution of the process goes through each of the stages and reaches a final state if no more stages have to be implemented. During this, the current stage of the logistics process can be conceived as its own progress status, whereby a set of stages are already completed and a set of stages have to be implemented. For instance, the goods of a transport from Dresden, Germany to Ashdod, Israel could just have reached the designated Italian port. Hence, different transportation legs of the process already have reached their final status, whereas the current leg is ongoing and the subsequent legs wait for their start. Technological speaking, the different stages a transportation leg has to go through construct a lifecycle of the leg.

The *Processing Engine* has the responsibility to keep track of these lifecycles and updates the internal lifecycle data accordingly. This enables users of the BCM to exactly see in which stage the observed transport is and react accordingly if, for example, delays occur. For this, the *Processing Engine* uses the *Collaboration Object Manager* to get access to the stored data and integrates current information (e.g., originating from the EPM) in the lifecycle of a certain CO. This integration is usually implemented by a stage transition within the lifecycle.

3.2.4. Collaboration Object Storage

As already described, the BCM integrates information from various sources and make it available for all involved stakeholders in a logistics process. For this, it is necessary that this information is persistently stored so that it can be loaded if a user demands it, even if the originating system is no longer connected.

The *Collaboration Object Storage* is the central data store of the BCM. It persistently stores the created instances of the CO-based data model and is able to select and load stored information. The storage is intended to be distributed. This means that the Collaboration Object Storage consists not only of one monolithic data base but each provider of the BCM can define his own data store. The distribution of the data avoids the central storage of possibly confidential information of all participants of a logistics process. A detailed functionality definition of this feature is not elaborated yet. We assume that the different data storages are connected via secure interfaces, which allows a controlled access and therewith, the collection of information that might be scattered over different concrete storage facilities. In addition to this, it is envision that the *Collaboration Object Storage* will work in a cloud-based manner and therewith, hide the concrete storage technology from potential clients. They just send storage and load requests and the underlying logic delegates these to the corresponding data storage. The cloud-based logic of the *Collaboration Object Storage* organizes the addressing of concrete data bases in order to achieve a persistent storage of all data.

4. Initial Technical Requirements Analysis

This chapter provides the initial analysis of technical requirements for the BCM. These requirements are used to identify the relevant technology areas, for which an initial State-of-the-Art analysis and preliminary assessment is presented in the subsequent chapters. For this, the chapter uses the BCM's basic building blocks, presented in Chapter 3, and determines requirements for each of them in a separated section. Based on a precedent explanation, particular requirements are identified and grouped into one of these categories: functional, technical or non-functional. The categorization is not final due to the fact that in such an early stage of the FInest project it is hard to determine which requirement addresses more a technical or only a functional aspect. During the refinement of the provided requirements a more accurate categorization based on exact definitions will be provided. Each requirement provides a label to be referenced in subsequent sections.

4.1. Requirements for the Security & Privacy Manager

Security and privacy management is a central requirement of the BCM. As pointed out in Section 3.2.1 the Security & Privacy Manager is concerned with it and restricts the access to information or checks them for integrity. However, the identification of users as well as third party systems (authentication) and the definition of allowed operations (authorization) is not in the scope of this component, and a component of the basic FInest platform is used. The task of the *Security & Privacy Manager* starts beyond basic user management duties. Based on the capabilities of an end-user or a third party systems the manager ensures that only allowed information is delivered. For this it is required that the *Security & Privacy Manager* defines the capabilities and disclosure level for the stored COs. In general, this encompasses three different aspects:

(i) Specification of the available attributes of the data model for particular end-users
(ii) Specification of special access rights to these attributes for particular end-users
(iii) Specification of visible parts of the lifecycle model for particular end-users

Apart from the mere definition of what data artifact can be accessed by which end-user or third party system, it is also required that the *Security & Privacy Manager* ensures that the restrictions are adhered.

Based on the previously presented explanation, the following technical, functional and non-functional requirements are identified:

Table 1 - Requirements for the Security & Privacy Manager

Functional Requirements		
Identifier	**Name**	**Description**
R101	Access Control	Implement an efficient and sufficiently precise access control mechanism. Define appropriate way to bind data artifacts to user and their roles.
R102	Restricted Access	Ensure that only authorized users and third party systems can access the stored data.
R103	Secure Transfer	Ensure secure and reliable information transfer at any time (e.g., through the Web, Future Internet Platform or Cloud-based system)
Technical Requirements		
Identifier	**Name**	**Description**
Rt101	Authorization	Implement mechanisms to restrict the access level and access rights of users and third party systems.
Rt102	Authentication	Provide mechanisms to validate user and third party systems identities.
Rt103	End-to-End Encryption	Use appropriate protocols to provide end-to-end encryption if data is delivered or received.
Non-Functional Requirements		
Identifier	**Name**	**Description**
Rn101	Efficiency	Applied security mechanisms must not harm the systems performance or lead to high system loads

4.2. Requirements for the Collaboration Object Manager

The central requirement for the *Collaboration Object Manager* is that it provides an appropriate modeling technology that is able to gather and reflect all (important) information of logistics processes. The current vision of the BCM calls this data model *Collaboration Object*, whereas this does not refer to any concrete modeling approach or technology. Essentially, a *Collaboration Object* can be any conceptual entity; a very specific artifact of a logistics process or a whole process. The term is just used in order to refer to a specific element in the internal data model of the BCM. For the implementation of these *Collaboration Objects* it is necessary that the used modeling approach considers the special demands of the T&L domain and gather all information that is related to a logistics process. Due to the variety of different T&L

scenarios make it will not be possible to integrate all end-user demands. The used modeling approach has to support this by providing the flexibility to adapt to concrete usage scenarios and customs demands.

T&L processes have a very close combination of data and lifecycle stages. The data describes the different attributes like characteristics of the transport goods, contract, or customs information. Different aspects of this data are needed in different stages of the lifecycle. For instance, a stakeholder could require the national transport by truck detailed information about the transport goods (width, height, weight, security classification, etc.) but no customs or contract information. The binding of information to explicit parts of the lifecycle introduce the requirement of an efficient combination of process information and general data.

Every logistics process is implemented within a highly instable environment, namely the global transport paths or network. A variety of situations can change the in-advance created transport plan. Traffic jams, flight controllers on strike or volcanic eruptions, which keep all planes on ground, are just a few illustrative examples for this. A potential modeling approach for the BCM has to support the changes that are caused by such unexpected conditions, by adapting the current model to the changed circumstances. Thus, a model of a logistics process cannot be fixed, but has to be modifiable at the runtime, which means during the execution of logistics process.

Directly related to the support for runtime changes is the requirement to keep track of this changes. If a logistics process cannot be run as previously planned this does not only affect the changed execution of the process. Previously planned transportation legs have to be canceled and this could entail cancellation fees. The information about this has to be preserved in the transportation model in order to support subsequent billing calculations or similar.

Like in every ongoing business, the requirement for constant changes of existing business processes exists also in the T&L domain. New customer demands have to be respected and new technologies could enable the implementation of complete new process. The selected modeling approach for the BCM has to support such changes and especially do not rely on “hard-coded” or fixed structures.

The previously described requirements target mainly on the used CO-based data model. In addition to this, it is of course necessary that the *Collaboration Object Manager* is able to create instances of COs described by the data model based information from external systems and manage these instances accordingly. This mean in detail, the persistent storage of created CO instances, the loading of information and the update of model instances if changes have been occurred.

Based on the previously presented explanation, the following technical, functional and non-functional requirements are identified:

Table 2 - Requirements for the Collaboration Object Manager

Functional Requirements		
Identifier	**Name**	**Description**
R201	CO-based Model	Define a model, based on *Collaboration Objects*, to describe logistics processes in a structured way
R202	Completeness	Gather all (necessary) information related to a logistics process
R203	Adaptability	Adaptation support for defined CO-based models at design time and at runtime (e.g., if certain transportation legs cannot be performed and the predefined plan has to be changed).
R204	CO Management	Load, store and update logistics COs from the underlying storage and provide facilities to transform externally provided information into a CO-based representation
R205	Standardized Data Formats	Currently, various standardized data formats are in charge in the T&L domain for data interchange. The most important of them have to be supported in order to enable communication of the BCM with existing legacy systems. Examples for this are EDIFACT or Cargo2000.
Technical Requirements		
Identifier		
Rt201	Modeling Approach	Choose a suitable modeling approach to define the CO-based data model (respect particular properties of T&L domain e.g., strong combination of data and process)
Rt202	Flexibe (Meta-) Model Definition	Constant changes in the logistics domain require constant changes on the defined CO-based data model. This has to be flexible to integrate these changes.

4.3. Requirements for the Processing Engine

The requirements for the *Processing Engine* partially overlap with the requirements for the *Collaboration Object Manager*. As described in the previous section, in transport and logistics processes exist a close relationship between process and data. The central requirement for the *Processing Engine* is that it is able to represent the current status of the transport process within the internal data model. Based on events delivered from systems outside the BCM (e.g., the *Event Processing Module* of the FInest module) it has to be able to update the current process status in order to keep the information within the CO up-to-date. At this point it is important to differentiate between the physical status of a logistics process in the real world and the representation in the internal CO-based model of the BCM. In order to make it clear, from now on, we use two different terms to indicate this difference: a **process stage** describes the real and physical situation of the logistics process; the **lifecycle** of a logistics process targets the representation within the internal CO-based model.

The requirements examination for the *Collaboration Object Manager* already pointed out that the T&L domain provides a wealth of different use case and scenarios. The *Processing Engine* has to be able to define and execute all possible implementations of lifecycles used in this domain.

The constant changes within the T&L domain also require mechanisms to facilitate the maintenance and adaptation of pre-defined logistics process stages, similar to the *Collaboration Object Manager*. Hard-coded or fixed definitions are not sufficient.

Based on the previously presented explanation, the following technical, functional and non-functional requirements are identified:

Table 3 - Requirements for the Processing Engine

Functional Requirements		
Identifier	**Name**	**Description**
R301	Process State Reflection	Reflect the current state of a logistics process in the CO-based model of the BCM
R302	Process Completeness	Define and execute every possible combination of process stages in T&L scenarios
R303	Process Change Support	Provide support for process changes at the design time of new logistics processes in order to integrate changes in the current transport process implementations.
Technical Requirements		
Identifier	**Name**	**Description**
Rt301	Process Definition and Execution	Implement feasible and reliable ways to define logistics processes. No “hard-coding” to support changes

4.4. Requirements for Collaboration Object Storage

The central task for the *Collaboration Object Storage* is to store instances of the CO-based model provided by the *Collaboration Object Manger*. Through querying technologies the selection and retrieval of desired information have to be possible. Modern business processes deal with high amounts of data volumes and used systems must be able to deal with these amounts in a scalable and efficient manner. The T&L does not mark an exception of this and the BCM also has to handle high amounts of data.

The variety of stakeholders involved in logistics processes raises the necessity for a precise selection, so that everybody is able to see the desired information. Due to the operation on a huge data base it is required that the *Collaboration Object Storage* is able to store and load information in an effective and scalable manner to do not harm the performance of the whole system.

The *Collaboration Object Storage* has to support distributed and cloud-enabled storage solutions. It has to provide logically central access to physically distributed data.

Additionally, different involved stakeholders probably have an interest that their confidential data is not managed by a data base out of their control or trust. For this reason, it is also required that the organization of the underlying storage infrastructure is visible to users so that they are able to determine where their probably confidential data is stored.

Based on the previously presented explanation, the following technical, functional and non-functional requirements are identified:

Table 4 - Requirements for the Collaboration Object Storage

Functional Requirements		
Identifier	**Name**	**Description**
R401	Performance and Scalability	Provide performance and scalability for large data amounts under consideration of the CO-based model structure (e.g., graph-based, relational, …)
R402	Integrated Security Management	Safe storage of obtained information by the provision of integrated security mechanism
R403	Distributed Storage	Implement the data storage in a distributed system. Enable the support for cloud-based storage.
R404	Transparency & Trust	Use of transparent and trustable storage infrastructure in order to enable user to determine how their probably confidential data is stored and managed.
R405	Suitable for CO-based Model	Native support for the used CO-based model (e.g., graph-based, relational, …)
Technical Requirements		

Identifier	Name	Description
Rt401	Encryption	Use encryption to safely store the gathered information, even in a distributed environment.
Rt402	Transport Security	Uses appropriate technologies to ensure the security and integrity of data within the distributed storage, especially during the data transfer
Rt403	Data Storage Technology	Use appropriate technologies that natively support the nature of the used CO-based model (graph-based, object-based, XML-encoded, …)
Non-Functional Requirements		
Identifier	**Name**	**Description**
Rn401	Reliability	Failure safe system
Rn402	Data Integrity	Ensure consistency and integrity of stored data
Rn403	Response Times	Provide sufficiently response and latency times in order to support real-world business scenarios.

5. Technical State-of-the-Art Analysis

We provide in this chapter a technical State-of-the-Art analysis for relevant technology areas. It assesses different technologies and approaches against their compliance to the defined requirements in Chapter 4. In the first section we will identify the relevant technology areas and subsequent to this, we give will a state-of-the-art analysis.

5.1. Selection of Relevant Technology Areas

In order to implement the vision of the BCM, relevant technology areas have to be found. To this end, we used the initial conceptual architecture in combination with the proposed capabilities of the BCM and identified five different areas. The following enlists these and describes why the certain area is of interest:

1. *Security and Privacy Management Technologies:* This area has to be examined in order to find suitable approaches to ensure the non-disclosure of confidential information within the BCM. The central point of usage of such technologies is the *Security & Privacy Manager*. However, a precise State-of-the-Art analysis within this area is already conducted in Deliverable D3.1 "Technological requirements and stat-of-the-art analysis" and therefore won't be discussed it in more detail here.
2. *Data and Process Modeling Approaches:* One of the BCM’s central capabilities is to provide end-users with a global end-to-end view on logistics processes. For this, the

representation of data and process information is a crucial point and appropriate approaches have to be identified.

3. *Cloud-based Hosting and Cloud-based Storage:* The whole FInest platform is supposed to use future internet technologies for the T&L domain. Cloud-based applications and data hosting is undoubted part of this technology stack. Thus, the BCM as one of the FInest core modulesIn the recent years the T&L domain has discovery the potential of cloud-based software to reduce costs and provide an easy to use personalized IT system. Even in 2009 [7] stated the important role of SaaS-based systems for logistics companies and nowadays, a growing number of solutions is available on the market. Only a few examples are: Rambase[2], Logistics Mall[3], CarrierNetOnline[4], logixcentral[5], OmPrompt[6] and e2open[7]. Hence, cloud-based systems play an important role for the future of logistics and FInest envisioned the implementation of all its modules in a cloud-based manner. Of course, the BCM as one of the FInest core modules, should also be implemented a cloud-based environment. For this reason it is necessary to investigate the State-of-the-Art in this technology area. However, the focus in this document lies on the transparent and trustable data storage in a cloud-based and therewith, distributed environment.

4. *User Interface Technologies:* In order to provide end-users of the BCM access to the managed information it is necessary to implement appropriate user interfaces. Hence, user interface technologies are relevant for the BCM. However, a discussion will be omitted due to the fact that it is already conduced in Deliverable D3.1.

5. *Services-based System Integration:* All modules of the FInest platform provide service-based interfaces to integrate other modules (this also includes user interfaces) or third party systems. The definition, implementation and provision of these interfaces is an important part and there exists a wealth of research with respect to this topic. However, the State-of-the-Art of technology area is already investigated by Deliverable D3.1 and therefore, will be omitted at this point.

In summary, the State-of-the-Art analysis for three of the five identified relevant technology areas is conducted in Deliverable D3.1. The subsequent sections will conduct a precise investigation of current research and development approaches for the remaining two areas.

5.2. Data and Process Modeling Approaches

In this section we investigate the State-of-the-Art for data and process modeling approaches. This reveals suitable approaches in order to implement the vision of the COs with respect to the previously identified functional and technical requirements in Chapter 4. Due to the wealth of different data and process modeling technologies, a first section will classify them and selects appropriate approaches. Subsequent sections will investigate the selected approaches in detail.

[2] http://www.hatteland.com/logistics.aspx?MenuItemId=304&Page=rambase
[3] http://www.ccl.fraunhofer.de/
[4] http://www.deltion.co.uk/
[5] http://www.logixcentral.com/
[6] http://www.omprompt.com/
[7] http://www.e2open.com/

5.2.1. Classification of the Modeling Approaches

The modeling of data structures is one key requirement since the very beginning of computer science. Due to the fact that especially business-related use cases are usually dealing with large datasets, the application of IT systems for these and the development of specialized programs facilitates the elaboration of different approaches to model data structures. One of the most important approaches until today is the Entity-Relationship-Model introduced by [8] in 1976. It provides mechanisms to express the relationship between data artifacts and is mainly used to describe data schemas for relational data base systems. Even modern products, such as Microsoft's Entity Data Model [9], are based on this approach, which is a huge indicator for its importance.

With the upcoming of object-oriented programming languages in the early 90ies, new modeling approaches became popular. The object-oriented modeling [10] describes data not by the means of entities and their relations but by classes, associations between them and class attributes. Standardized modeling languages, like the Unified Modeling Language (UML), or modern frameworks, like the Eclipse Modeling Framework (EMF), had a huge impact on the success of this modeling approach, so that it is today widely accepted, especially in domains closely related to application development.

Other data modeling approaches were introduced during the upcoming of generic data exchange formats like XML or JSON. The majority of them uses their own schema definition language in order to define the structure of instance documents and therewith, the used data model. Other schema languages have been introduced by the developments around the Semantic Web movement and the used data exchange formats, for example RDF and RDF/S.

However, all these different approaches are focusing on the definition of data structures and this is only one aspect needed for the implementation of the envisioned COs. The more important part is the creation of a process model for logistics processes so that the current stage of it can be represented and previous as well as subsequent stages are derivable. For this, a mere data modeling is not sufficient [11] and due to this a detailed investigation of these approaches will be omitted here. It could be fetched in subsequent deliverables, which will provide a more comprehensive examination for the implementation of COs, along with a consideration of the internal data representation.

Thus, the focus of this section lies on approaches that are able to define business processes and three different of theses are identified: process-centric modeling, data-centric modeling, ontology-based process modeling. The last one provides a process definition based on established ontology languages and semantic frameworks [12]. Existing systems can be supported by an automated transformation in other process models and languages. However, it seems that this approach suffers the same drawbacks as semantic descriptions in general and Semantic Web Services in particular. The creation of appropriate, ontology-based description models is a very elaborate task and due to this their impact on the industrial domain is very low. For this reason, a detailed investigation of this approach is also omitted here.

Consequently, the process-centric and the data-centric modeling are described in the subsequent sections. These will provide a State-of-the-Art analysis for research and development efforts and present current implementation technologies.

5.2.2. Process-centric Modeling

To date the predominant technologies to organize and manage business processes are undoubtedly based on a process-centric approach [6], [13]. This approach suggests the process representations as a set of entities upon which a set of actions can be performed. Edges between the entities associate these with each other and establish the process description, similar to the notion of Finite State Machines (FSM). Entities have attached attributes (data) but are passive and do not manipulate these directly. Actions are invoked on a set of entities affect the attached attributes and may change their internal status. By this, the process-centric modeling approaches reflect the current progress of real-world business processes. However, as the name of the approach suggests the focus lies on the description of a process model and the attached data or attributes are usually handled separately.

An alternative and more common term for the process-centric modeling of business processes is workflow management. A system that defines, create and execute such a workflow is denoted as Workflow Management System (WMS) and [14] defines it as follows:

> *A system that defines, creates and manages the execution of workflows through the use of software, running on one or more workflow engines, which is able to interpret the process definition, interact with workflow participants and, where required, invoke the use of IT tools and applications.*

Of special interest for the use within the BCM are the supported process definition languages. In accordance with requirement R302 and R303 it is necessary that process definitions can be created that reflect any possible scenario within the T&L domain. Additionally, these definitions have to be easily maintainable in order to support process changes. In general, four commonly used process definition languages can be distinguished:

- Petri nets
- Event-driven process chains
- Activity Diagrams / State machines
- Business Process Model and Notation (BPMN)

Usually one of the approaches is supported by a concrete WMS implementation (to date in most cases BPMN). Naturally, there are also proprietary languages in charge which target on a particular use case or domain.

5.2.2.1. Existing Tool Support

A wealth of different WMS implementations is currently available for nearly all possible demands and requirements. Of course, not all of these systems target business processes and process-centric modeling approaches can also be used for other tasks (e.g., technical process definitions like service composition or choreography). However, the majority of the available tools support BPMN which emphasizes the predominate role of process-centric modeling approach for the definition of business processes. A BPMN-based modeling and execution facility is nowadays usually included for modern business IT systems. The SAP's Netweaver platform [15] or IBM's WebSphere [16] are two of the most prominent examples. In addition to this are a variety of Open Source Tools available. In [17], different WMS for Java based platform are listed (without the claim of completeness). It can be assumed that other programming platform provide a comparable tools set in this area. Apache ODE [18] is one of the most prominent open source tools. Although it uses WS-BPEL for the process definition, it is still capable to define business processes. Its primary target on web services executive as executive parts of business process makes it primarily interesting for service-enabled business environments. Another more lightweight approach is Activiti [19], which tries to address developers, system administrators as well as business people. The provisioning of sophisticated editor toolkits facilitates the creation of workflow processes. A more detailed investigation of the available tool landscape is omitted here due to the fact that the overall approach is focused.

5.2.2.2. Runtime Change Support

All introduced systems do not support changes at runtime. Instances of a predefined process description are created and executed. A variation during the runtime is not supported. However, as described in Section 4.3 the T&L has to deal with a lot of unexpected changes during the execution of a logistics process. This has led to the requirement R203 which demands support for the integration of runtime changes.

One system that supports adaptability of traditional workflows is ADEPT [20] . It allows a hierarchical process definition and uses process schemas for this. This allows the user to add or delete process elements as well as changing their execution order during runtime. However, this is only supported in exceptional cases and the approach is built on top of the own simple workflow specification by using algorithms and theorems based on Petri-nets. We believe that a more native support for runtime changes is highly desirable because this approach introduces additional complexity and could be cumbersome.

Another approach is provided by Worklets and is presented in [21]. A Worklet is a part of a process definition that can be assembled together in order to define the whole workflow model. How different Worklets can be combined is defined by template processes. These templates provide a framework and define hook points where different Worklets can be integrated. The selection of a Worklet from the repertoire is based on rules. Both, rule and Worklets, can be added during runtime and thereby, provide rich capabilities for reuse and runtime variability. However, the main disadvantage is the hook points are defined in the template and cannot be defined dynamically.

In [22], a approach is described to enable process variability by the use of state machine to explicitly model the behavior of every process step. Thus, this stage machine defines a second abstraction level for the definition of workflow models, which can be beneficial for non-technical users. Process steps can be composed without knowledge of their internal model. Only for the support of runtime changes is state machine is used. Every process step can be halted if it is in a consistent state. After stopping a process step it can be modified or exchanged as long the output of the activity remains the same due to the fact the subsequent step depend in it.

5.2.2.3. Applicability for Logistics Domain

Despite the various approaches to support runtime changes of workflow definitions it seems that most of them are difficult to handle and do not allow an easy integration of occurred changes. The initial vision of process-centric modeling rather focuses on the execution of a predefined workflow. Due to this the presented approaches just defines another abstraction level to support runtime changes and this do not provide the envisioned easy support for runtime adaptability (cf. requirement R203).

Even though workflow management systems would support this process-centric modeling approaches share a common disadvantage. As mentioned before, they focus on the definition of control flows. By this, only the current stage of the process is relevant and data attached to the whole process (not to particular stages) is moved to the background. This problem is described by [23] as *context tunneling*. Instead, providing a user with all necessary information about what he or she has to do, process-centric modeling defines rules how things should be done. This could harm the understanding of the overall goals and hampers active solution management. However, this is necessary in order to solve the complex problems occurring in the T&L domain.

5.2.3. Data-Centric Modeling

In order to solve the problem of context tunneling other research approaches emphasizes more the data aspect of business processes. This has led to so-called data-centric modeling approaches, whereby they treat both data and processes as first-class citizens. This is a fairly new research area and the first (concrete) approach was elaborated in 2003 by [6]. It is based on the idea of promoting all aspects a business operates on, as artifacts. An artifact could be any "*concrete, identifiable, self-describing chunk of information that can be used by a business person to actually run the business*" [6]. Examples are sales orders, customer profiles or transportation legs. Due to the fact that *"artifact"* is already a well established term in BPMN, the approach was recently renamed to *"Entity-centric modeling"*, so that the original name can only be found in older publications. We will use this new term from now on and denote the modeling elements as *entities* (instead of artifacts).

The main difference to the process-centric modeling is the combination of data (*information model*) and process (*lifecycle model*) aspects of a business process in a single construct. Additionally, every entity provides information about applicable tasks and thereby, defines rules for state transitions of the internal process model. By this, entities serve as basic building blocks from which business operations and process models are constructed. The main advantage of the approach is that it introduces different levels of abstraction and enables a natural modularity as well as componentization of business processes [24].

5.2.3.1. The Business Entity Methodology

In this section we present an example methodology for data-centric modeling approaches. We chose the entity-centric modeling approach due to the fact it gained a lot of interest from the research community during the last years. Besides this, also other modeling methodologies use the data-centric approach. An example for this is the case handling paradigm, as presented in [23]. This approach organizes all work that has to be done around cases (or case files). Each of these contains all the relevant data to complete the task but without any work or control flow information. The implementing party of the task is completely free to define their own business logics in order to reach the process goals. However, we decided to investigate the entity-centric approach more in detail due to the fact that the complete freedom of data manipulation seems not to be suitable for the T&L domain, where processes have to conducted in a certain order and accordingly a particular schema.

The inventor and main contributor to the research area of entity-centric modeling is IBM Research. After the first vision of business entities (cf. [6]) they constantly evolve the approach and applied it in internal and external projects [25], [26]. This revealed a set of additional requirements which are necessary for an application in a productive environment. Based on this the key elements of the entity-centric mindset are defined.

In contrast to the uni-dimensional process-centric approach, the entity-centric approach provides *"four explicit, inter-related but separable dimensions in the specification of a business process"* [13]. Figure 5 illustrates these and the following provides a detailed explanation.

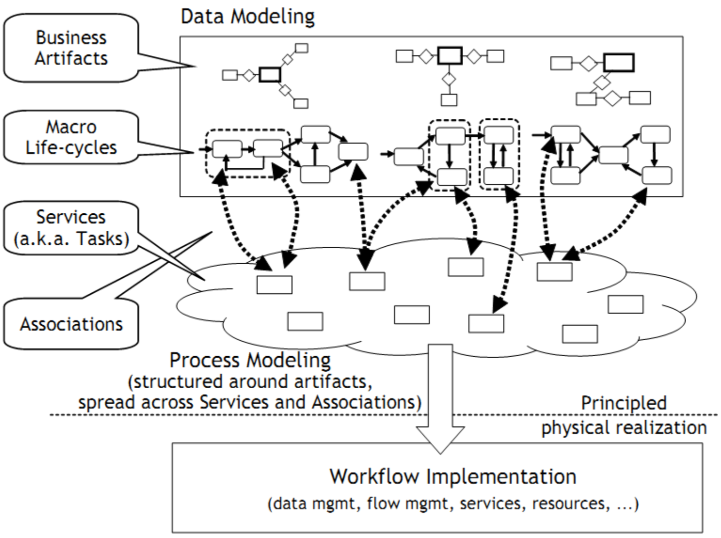

Figure 5 - Dimensions of entity-centric modelling [13]

Business Artifacts. The name of this dimension could be a little bit confusing due to the fact that [13] focuses with the term 'artifact' at the data part of a business entity. Hence, this dimension addresses the information model, which shall incorporate all information that is needed to capture business process goals and allow the evaluation how these can be achieved. Additionally, this model may include metadata such as a unique identifier, relations to other entities, the change history of the entity and information about the current state. There are no restrictions regarding the implementation of this models and every suitable technology can be used. The paper presented in [6] suggests key-value-pairs as well as XML documents or more sophisticated approaches, like databases.

(Macro-Level) Lifecycle. The second major part of an entity is the lifecycle. In consists of a set of stages through which the entity progresses during the execution of a business process. The lifecycle is directly related to the information model of an entity as indicated by the box in Figure 5. This lifecycle consist of key and business-relevant stages in the potential evolution of an artifact (e.g., from inception to disposition and achieving), usually identified business domain experts. The integration of data and the lifecycle description is one of the main benefits of the

entity-centric modeling approach. As for the information model, there a no restrictions defined for the implementation technology for the lifecycle.

Due to the fact that Figure 5 has some deficits in the depiction of the information models and lifecycle models within business entities, Figure 6 shows an example of a fictive business entity ‘Job Application’. The upper part of the figure contains the lifecycle model (denoted as *Marco-Level Lifecycle* in Figure 5) of the entity. It consists of all stages through which the entity can evolve during an application process. The part below shows the information model (denoted as *Business Artifact* in Figure 5) of the entity, which is directly associated with the lifecycle model. It is important to understand that data and process information build a mutual combination in the entity-centric modeling approach.

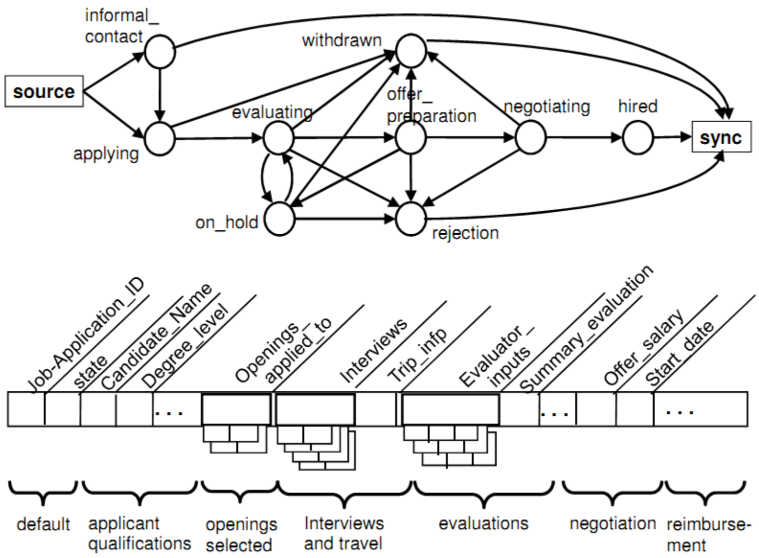

Figure 6 - Example Business Entity 'Job Application' [27]

Services (Tasks). The lifecycle of a business entity contains little or none information about how and why lifecycle stages are changed. A service encapsulates a unit of work within the business process, which has to be meaningful at least in two aspects: (i) it has to be a step towards the achievement of the business goals; (ii) the segmentation of the business process in services should facilitate administrative and organizational structures. Figure 5 uses the term ‘service’ to emphasize the close correspondence of a task that can be invoked on an entity and services of a service-based environment. The execution services may affect the lifecycle model of a certain business entity and trigger a stage transition.

Associations. This dimension describes under which constraints tasks can be invoked and links them to lifecycle stages. Associations can either be defined in an imperative or declarative manner. The first, for example, can use state-machines to define which tasks are executed on which transitions. The latter, uses a set rules to describe the constraints of a specific task and this can be accomplished by the use of logic programming for instance.

5.2.3.2. Existing Tool Support

Due to the fact that the entity-centric modeling approach is a relatively new research area, nowadays only a few tools are available. The most elaborated tool support is provided by IBM the main contributor to this research area. The Siena system [28] provides a user interface to design entity-centric workflows, which can be used to describe the interoperation of web services. Another tool, which supports this type of business process modeling, is called BELA [29]. It is also developed by IBM and is based on the WebSphere product line. Apart from this tools, in 2010 started a publicly funded project, called ACSI (Artifact-Centric Service Interoperation) [30], targeting on (web) service interoperation by the use of the entity-centric modeling approach. For this, the project also develops a runtime engine for business entities that also can be used in other contexts. In accordance to the Description of Work, the project plans to release parts of the developed solution under open source licenses. Unfortunately, there is no more precise information available about this.

In comparison with the tool set available for workflow modeling, business-entity-based modeling is only supported by a very few programs. However even though existing tool support is desirable, it is not a direct requirement to achieve an affordable implementation of the data-centric modeling approach. Existing tools and technologies can be combined in order to realize the vision. For example, as we have described in Section 5.2.2.1, exists a wealth of different workflow management systems. These can be used to implement the lifecycle model of an entity and in combination with data modeling technologies (XML, object-oriented approaches like EMF, etc.) it is feasible to implement the information model. A rather lightweight engine just has to provide glue code to bind these technologies together. An example of such an implementation is provided by [22]. However, the inventors of the entity-centric modeling approach do not recommend this implementation style, due to the fact that the process description based on traditional workflow languages is cumbersome.

5.2.3.3. Applicability for Logistics Domain

If a business process is modeled with an entity-centric approach, it consists of a set of business entities, whereby each addresses a certain aspect of the process. Entities store connections to related entities and establish with this a graph that represents the whole business process. The encapsulation of the data and behavior in one entity provides many benefits for a flexible runtime changes. For example, if a change in the process has to be integrated, it is sufficient to remove the corresponding entity from the graph and substitute it with one or more replacements entities, or if a transport leg cannot be implemented it is sufficient to exchange the corresponding business entity. New business entities which describe the replacement transport legs can be integrated just by updating the references of the original entities of the process.

Security and Privacy is a key aspect of the BCM (cf. requirements R401 – R405). The present data-centric modeling approach based on business entity provides native technologies to facilitate the introduction of appropriate mechanisms for this. In [31], the authors describe the introduction of different ‘views’ for particular business entities. Each view can collapse or condense different stages of the internal lifecycle model to one stage, so that details can be hidden from clients. Views are also able to restrict the access to the internal data model and therewith, are an excellent foundation to implement powerful privacy mechanisms. In addition to this, the paper also describes 'windows' as a customized perspective of the whole business process model. A window is able to hide certain business entities from the client and is of special interest if the whole process contains different confidential parts. Access rights to business entities can be described by providing Create-Read-Update-Delete-Append-Execute (CRUDAE) permissions for every business entity and for parts of the internal data and lifecycle model. These sophisticated security and privacy mechanisms can be integrated in the entity-centric process model. We consider this as a important benefit due to the fact that these issues are already addressed at the modeling level and are not just part of a complete independent and higher abstraction layer.

Another huge benefit is the natural combination of process (behavior) information and data of business entities. The requirement Rt201 describes the important role of this for T&L domain. The presented data-centric modeling approach is able to implement this in very elegant manner.

5.3. Cloud-hosting Technologies and Data Storage Approaches

In this section we describe the current State-of-the-Art of existing cloud hosting and data storage solutions. Our aim is to reveal approaches that enable a transparent and trustable data handling, so that users of the BCM are able to determine how their potential confidential data is stored and managed. These are central requirements of BCM for its cloud-based hosting (cf. Requirements R403 and R403) and we consider this a major feature, which has to be supported. For this, we will briefly introduce the concept of cloud-computing with its different manifestations, introduce different kinds of cloud implementations and present projects as well as research efforts that consider trustable approaches.

Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g. networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.

5.3.1. Layers of Cloud Computing

Accordingly to Furht and Escalante [32] *"cloud computing can be viewed as a collection of services, which can be presented as a layered cloud computing architecture"*. They are describing the five general layers of cloud computing architectures [32]:

1. **Data Storages as a Service (dSaaS)** includes servers and data storages available remotely which serving as storage for all kinds of data including application data.
2. **Virtualization** stands for virtualized servers running in the cloud which results in fewer servers needed that in non-cloud environments because one physical server can host several virtual servers.
3. **Infrastructure as a Service (IaaS)** stands for computing resources which are accessible as services.

4. **Platform as a Service (PaaS)** uses IaaS, adds an operating system and a required software stack to run a specific application.
5. **Software as a Service (SaaS)** includes applications which are available as a service which means that they will run remotely from the cloud without any application specific files installed locally.

To run an application from the domain-specific extension of the FI PPP Platform all layers are necessary.

The FInest platform will build upon the PaaS layer and provides its software solutions as SaaS to its end-user. However, the different FInest modules are not intended to provide all the layers below but use existing technologies and frameworks. For this reason it is necessary to identify potential candidates in the layers 1 – 4.

5.3.2. Key Technologies and Principles for Cloud Hosting

The following concepts and technologies are crucial for the success of a cloud implementation [32]:

- **IaaS** as already mentioned stands for physical resources (mainly network switches, servers and data storages) which are not sold but their usage is offered as storage and computing resources.
- **Server Template Technology** means that preconfigured server templates (images) are available from which the customer can choose at setup which accelerates the web application deployment. A template contains an operating system and a software stack.
- **Virtualization** is mainly important for the cloud provider because it enables the sharing of physical servers among several virtual servers.
- **Dynamic Orchestration of Resources** means that additional resources will be assigned to applications if needed. As a result, cloud applications can always have enough computing power and storage to perform well.

5.3.3. Cloud Privacy

A cloud can either be hosted **public**, **private** or as a **hybrid** of both approaches. In contrast to a public cloud, where data can be held anywhere, organizations can built private clouds with own IT resources or order one from a cloud provider to have full control of the stored data. Additionally, hybrid clouds can be formed using public and private IT resources which may be useful if only a part of the stored data should be threaded absolutely confidential and the other parts are not mission-critical.

For the FI PPP Core Platform and its domain-specific extensions like the one planned in FInest public clouds may **not** be an option because T&L companies, like all companies, may own a lot of mission-critical data. Therefore, transparency is needed (Requirement R404) which means that the company always knows where its data is physically located. As a result, only private or hybrid clouds may be used.

The main benefits of public clouds are the high scalability and the payment by services and capacities as needed.

5.3.4. Cloud Computing Platforms

In this section main cloud computing Platforms are presented which also focus on transparency. This includes platforms provided by IBM, Amazon, Microsoft and Google which are the main players in cloud hosting. In general, all cloud computing platforms work on the same basic principles (section **Error! Reference source not found.**); however, in this section it will mainly be focused on transparency and trust of the provided infrastructures.

5.3.4.1. IBM SmartCloud

IBM offers one of the brightest ranges of implementation options for their clouds which enable customers to customize their cloud specific to their needs. Additionally, IBM enables customers to save their server configurations as standard-images in a private repository.

In terms of cloud security IBM provides a security framework which includes a concept on which IBM is responsible for correct authentication and authorization while the company takes care of the cloud user management. Additionally, IBM ensures that the IT-environment, which includes the networks, server and end devices, is safe of unauthorized access, damage and misuse [33].

5.3.4.2. Amazon Elastic Compute Cloud (Amazon EC2)

Amazon does not follow a special transparency strategy in terms of letting the customers know where their data is located. Amazon itself says about the physical location of their datacenters: "AWS datacenters are housed in nondescript facilities" ([34], p.5); therefore it is not reproducible for a customer where his data is located. However, Amazon also has a complex security program with lot of security techniques to ensure the customer data security [34].

5.3.4.3. Google App Engine

Google's App Engine seems to be more easy to use by developers to develop and deploy simple applications. Google cares completely about allocation without any customer interactions. However, no details about security mechanisms used by Google's cloud have been found. The company justifies its security level only by referring to Google's 10 years of experience in hosting web applications [35] and the achievement of three certificates for data protection and security – *SAS 70 Type II*, *SSAE 16 Type II* and *ISAE 3402 Type II* [36]. However, also major downsides for organizations are known if they host their applications at Google's infrastructure, as [37] shows.

5.3.4.4. Microsoft Azure

Microsoft's cloud platform Azure is secured via a bright range of security technologies. Beyond the authenticating access to data also isolation of components from others is a main aspect Microsoft focuses on. For example each virtual machine has several access levels (for administrator, guests and others) which are isolated from each other. Also the different virtual local area network (VLAN) instances are totally isolated from each others, as well as, the customer access to the infrastructure from the customer access to the application and data storages. Azure uses SSL encryption for data transfer and different encryption algorithms to encrypt data in storages [38].

5.3.5. Research Projects on Cloud computing

In this section we briefly introduce two research projects which target on the introduction of trustable cloud-infrastructures and private clouds. These research efforts can be used to address unresolved issues in established cloud solutions, as presented in the previous section, with respect to the identified requirements for the BCM.

5.3.5.1. TClouds

The FP7 Project TClouds focuses on privacy protection in cross-border infrastructures as well as on ensuring resilience against failures and attacks, but *"without limiting the solution to just a physically separated private cloud"* [39]. The aim is privacy *"without losing the advantages regarding cost-efficiency, scalability and data availability"* [39].

One of the key ideas is Twin Clouds which consist of the usage of two clouds – a trusted cloud and a commodity cloud (public cloud). The clients are only connected via secure connections to the trusted cloud which executes the security-critical operations while the commodity cloud executes the performance-critical operations because more IT resources will be assigned to it. All data which will be sent to the commodity cloud will be encrypted before by the trusted cloud and then sent via a high bandwidth connection [40]. Figure 7 shows the concept again.

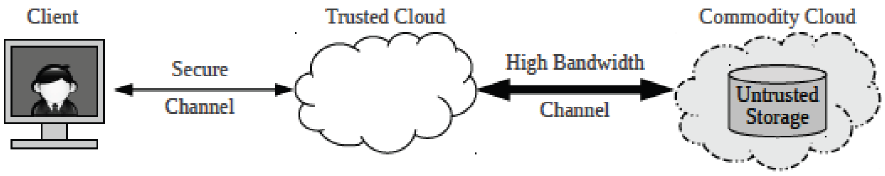

Figure 7 - Twin Clouds model with Client, Trusted Cloud, and Commodity Cloud [40]

Additionally, the project focuses on privacy functionality in cross-cloud communication protocols, new open security standards and components for effective cloud security management, among others [40].

5.3.5.2. RESERVOIR Project

The RESERVOIR is an EU-project of the FP7 Framework and provided the RESERVOIR Framework which enables users for business and government to build their own cloud on-demand. The framework enables non-IT-specialists to benefit from all cloud computing advantages like reducing investment, reducing operational costs and increasing energy efficiency, for instance. Additionally, in RESERVOIR several management tools have been build to support the cloud creation [41]. This framework could be used by organizations to build their own private cloud for their instance of the FInest platform extension.

6. Initial Identification of Potential Generic Enablers

The Chapter 5 of Deliverable D3.1 ("Initial Generic Enablers Requests") provides a complete list of requests for Generic Enables from all FInest work packages. Due to this, this chapter here refers to this section in D3.1 and omit a precise description of Generic Enablers for the BCM. The complete list also can be accessed through the following Web site: http://www.finest-ppp.eu/index.php/project-results/generic-enablers.

7. Selection of Technology Baseline

This chapter provides a preliminary selection of the technology baseline to implement the presented vision of the BCM. This is the basis for subsequent investigations of concrete candidate technologies in the following deliverables. The presented technical state-of-the-art analysis conducted in Chapter 4 and the investigation of potential Generic Enables from the FIWARE platform (cf. Chapter 6) is the foundation for this selection.

In general, wherever it is possible the usage of a Generic Enabler will be preferred instead of other technologies. This means, only if no Generic Enabler that is able to fulfill the defined technical, functional and non-functional requirements for a certain component of the BCM, an 'external' technology will be used.

The 'Selection of Relevant Technology Areas' in Section 5.1 identified five different areas of interest for the potential technologies. A state-of-the-art analysis for three of them is already conducted in Deliverable D3.1 due to the fact that these are also relevant for other modules of the FInest platform. In order to be compliant and interoperable, the selection of the technology baseline of the BCM for these areas is accordingly to Deliverable D3.1 and hence, a consideration will be omitted here.

7.1. Technology Baseline for Data and Process Modeling

In Section 5.2 we showed that entity-centric modeling approaches provide many benefits for the modeling and representation of logistics process in comparison to traditional process-centric approaches. The encapsulation of related data in business entities, the mutual combination of data and process information as well as the direct support for runtime makes an application of this approach hugely desirable. It is able to fulfill all defined requirements in Section 4.2 and therefore, it is chosen as used modeling approach for logistics process in the BCM.

However, the implementation of the modeling approach is still an open question. There is no Generic Enabler for entity-centric modeling available and due to the unique demands it can be assumed that there also will none provided in the future. In addition to this, the chosen modeling approach is a rather new research area and not well-supported by existing tools. Although, there are research results available which describe the implementation of an entity-centric modeling approach by the reuse of well established workflow engines, this kind of implementation seems very limited? . Thus, it is necessary to identify a feasible approach to implement the vision of entity-centric modeling for logistics processes and further efforts have to identify appropriate implementation technologies.

7.2. Technology Baseline for Cloud-hosting Technologies and Approaches

We described in Section 5.3 currently available cloud hosting and storage solutions of the most important players in this area. We assumed that each of these 'big players' is able to provide sufficient performance and scalability, even for huge data sets and many involved users. Thus, we focus more on the provided security, transparency and trust mechanism. Our investigation showed that the investigated cloud-hosting providers use different security policies and provide very different detailed information about this. Google, for example, just refers to their 10 years experience for web hosting solutions and provide no further information about used security mechanisms. All providers only give vague information about the used infrastructure (especially, Amazon), which drastically hampers the transparency and trust. For this reason we do not recommend the use of these systems to host an instance of the BCM and hence, we will not use as potential target platforms in the future development process.

A more promising approach is the use of the cloud-hosting facilities of FIWARE to implement private or hybrid cloud solutions. A private cloud is only accessible for a defined set of users. This could provide the necessary trust to encourage users of the BCM to share their data by the use of used infrastructure. However, the Generic Enablers of FIWARE, which addressing the cloud hosting, do not provide sufficient support for transparency so far. Users are currently not able to determine the setup of the underlying infrastructure and hence, could not determine how their data is handled. For this reason it is necessary to introduce additional capabilities into the FIWARE platform that enables user to get desired transparency information. The TClouds project provides auspicious approaches for this problem.

Hence, we define the preliminary technology base line for *Cloud-hosting Technologies and Approaches* as follows: The basic building block is provided by the Cloud Hosting Generic Enablers (especially, the *Object Storage* GE) of the FIWARE platform. Occurring transparency and trust issues are addressed by the introduction of further approaches, as provided by the TClouds project for example.

8. Conclusion

In this document we provide the initial requirements analysis and a preliminary selection of the technology baseline for the BCM of the FInest platform. After an introduction into the general vision of the BCM in Chapter 1, we conducted investigation of the as-is situation for collaboration systems used in the T&L domain today. In Chapter 2, we presented the initial conceptual design of the BCM, which builds the basis for the subsequent preliminary technical and functional requirements identification in Chapter 4. In Chapter 5 we used the previously defined requirements and the conceptual architecture to conduct a State-of-the-Art analysis in order to find applicable R&D results. Based on this, we selected an initial technology baseline for to implement the BCM (cf. Chapter 7). To ensure the alignment to the FI PPP platform, we identified potential Generic Enables in Chapter 6.

With the progression of the project and the gathering of more domain knowledge as well as partner feedback, the identified requirements will be refined. In addition to this, the next steps also include a more detailed specification of the conceptual architecture and based on this a concrete selection of technologies. Very crucial point in this regard is the identification of an appropriate and suitable implementation approach for the envisioned *Collaboration Objects* based on an entity-centric modeling approach.

9. References

[1] A. Hartwig and T. Rinza, “Internet basierte Supply Chain Collaboration in modernen Logistiknetzwerken,” 2002. [Online]. Available: http://logistics.de/downloads/84/2d/i_file_45268/ Internetbasierte_SC-Cooperation.pdf. [Accessed: 09-Sep-2011].

[2] S. Kummer, H.-J. Schramm, and I. Sudy, *Internationales Transport- und Logistikmanagement*, 2nd ed. Facultus Wien wuv., 2009.

[3] IATA Cargo 2000, “Cargo 2000,” 2011. [Online]. Available: http://www.iata.org/whatwedo/cargo/cargo2000/Pages/index.aspx, last access on 9/9/11. [Accessed: 09-Sep-2011].

[4] IATA Cargo 2000, “Purpose and Benefits of Cargo 2000.” IATA Cargo 2000, 2011.

[5] IATA Cargo 2000, “Master Operating,” 2011. [Online]. Available: http://www.iata.org/whatwedo/cargo/cargo2000/Pages/master-operating-plan.aspx, last access on 9/9/11. [Accessed: 09-Sep-2011].

[6] A. Nigam and N. S. Caswell, “Business artifacts: An approach to operational specification,” *IBM Systems Journal*, vol. 42, no. 3, pp. 428-445, 2003.

[7] A. Mesher, “SaaS 2.0: How the Logistics Industry is Driving an Evolution in Software-as-a-Service,” 2009. [Online]. Available: http://logisticsviewpoints.com/2009/09/22/saas-2-0-how-the-logistics-industry-is-driving-an-evolution-in-software-as-a-service/. [Accessed: 23-Sep-2011].

[8] P. Pin-shan, “The Entity-Relationship Unified View of Data Model-Toward a,” *Database*, vol. 1, no. 1, pp. 9-36, 1976.

[9] J. A. Blakeley, D. Campbell, S. Muralidhar, and A. Nori, “The ADO.NET entity framework,” *ACM SIGMOD Record*, vol. 35, no. 4, p. 32, Dec. 2006.

[10] James Rumbaugh, J. Rumbaugh, and F. Eddy, *Object-Oriented Modeling and Design*. Prentice-Hall, 1991.

[11] B. Curtis, M. I. Kellner, and J. Over, “Process modeling,” *Communications of the ACM*, vol. 35, no. 9, pp. 75-90, Sep. 1992.

[12] A. Koschmider and A. Oberweis, “Ontology based business process description,” in *Proceedings of the CAiSE*, 2005, vol. 5, pp. 321–333.

[13] R. Hull, “Artifact-centric business process models: Brief survey of research results and challenges,” *On the Move to Meaningful Internet Systems: OTM 2008*, pp. 1152–1163, 2008.

[14] Workflow Management Coalition, *Workflow Management Coalition Terminology & Glossary*. 1999.

[15] SAP AG, “Components & Tools of SAP NetWeaver - SAP NetWeaver Business Process Management,” 2011. [Online]. Available: http://www.sap.com/platform/netweaver/components/sapnetweaverbpm/index.epx. [Accessed: 13-Sep-2011].

[16] IBM Inc., “WebSphere MQ Workflow,” 2011. [Online]. Available: http://www-01.ibm.com/software/integration/wmqwf/flowcoal.html. [Accessed: 13-Sep-2011].

[17] Java-source.net, “Open Source Workflow Engines in Java,” 2004. [Online]. Available: http://java-source.net/open-source/workflow-engines. [Accessed: 13-Sep-2011].

[18] Apache Foundation, “Apache ODE,” 2011. [Online]. Available: http://ode.apache.org/. [Accessed: 13-Sep-2011].

[19] Activiti Group, “Activiti,” 2011. [Online]. Available: http://www.activiti.org/. [Accessed: 13-Sep-2011].

[20] C. Hensinger, M. Reichert, T. Bauer, T. Strzeletz, and P. Dadam, “ADEPTworkflow-advanced workflow technology for the efficient support of adaptive, enterprise-wide processes,” *Proc. Software Demonstration Track, EDBT*, pp. 29–30, 2000.

[21] M. Adams, A. ter Hofstede, D. Edmond, and W. van der Aalst, “Worklets: A service-oriented implementation of dynamic flexibility in workflows,” in *On the Move to Meaningful Internet Systems 2006: CoopIS, DOA, GADA, and ODBASE*, R. Meersman and Z. Tari, Eds. Springer Berlin / Heidelberg, 2006, pp. 291-308.

[22] K. Erler, “Design and prototypical Implementation of an Entity-centric Collaboration Environment,” Dresden University of Technology, 2011.

[23] W. Vanderaalst, M. Weske, and D. Grunbauer, “Case handling: a new paradigm for business process support,” *Data & Knowledge Engineering*, vol. 53, no. 2, pp. 129-162, May. 2005.

[24] D. Cohn and R. Hull, “Business artifacts: A data-centric approach to modeling business operations and processes,” *Bulletin of the IEEE Computer Society*, vol. 32, no. 3, pp. 1-7, 2009.

[25] K. Bhattacharya and R. Hull, “A data-centric design methodology for business processes,” *Handbook of Research on Business*, pp. 1-28, 2009.

[26] K. Bhattacharya, N. S. Caswell, S. Kumaran, A. Nigam, and F. Y. Wu, “Artifact-centered operational modeling: Lessons from customer engagements,” *IBM Systems Journal*, vol. 46, no. 4, pp. 703-721, 2007.

[27] ACSO Project, “ACSI Description of Work (Part B),” *Work*. ACSI Project, 2010.

[28] D. Cohn, P. Dhoolia, F. Heath, F. Pinel, and J. Vergo, “Siena: From powerpoint to web app in 5 minutes,” in *Service-Oriented Computing – ICSOC 2008*, 2008, pp. 722-723.

[29] K. Bhattacharya et al., “A model-driven approach to industrializing discovery processes in pharmaceutical research,” *IBM Systems Journal*, vol. 44, no. 1, pp. 145–162, 2005.

[30] A. Project, “ACSI Project Homepage,” 2010. [Online]. Available: http://www.acsi-project.eu/. [Accessed: 13-Sep-2011].

[31] R. Hull and N. Narendra, “Facilitating workflow interoperation using artifact-centric hubs,” *Service-Oriented Computing*, 2009.

[32] B. Furth and A. Escalate, *Handbook of Cloud Computing*, 1st ed. Springer, 2010.

[33] IBM Corporation, “Auf Unternehmenabgestimmte Infrastrukturvirtueller Maschinen auf der IBMCloud.” 2011.

[34] Amazon Inc., “Amazon Web Services: Overview of Security Processes.” 2011.

[35] Google Inc., “Why App Engine,” 2011. [Online]. Available: http://code.google.com/appengine/whyappengine.html. [Accessed: 15-Sep-2011].

[36] E. Feigenbaum, “Security First: Google Apps and Google App Engine complete SSAE-16 audit,” 2011. [Online]. Available:

http://googleenterprise.blogspot.com/2011/08/security-first-google-apps-and-google.html. [Accessed: 15-Sep-2011].

[37] G. Rogers, "The problem with Google Apps Engine," 2008. [Online]. Available: http://www.zdnet.com/blog/google/the-problem-with-google-apps-engine/1002. [Accessed: 15-Sep-2011].

[38] Microsoft Corporation, "Windows Azure Security Overview," 2011. [Online]. Available: http://www.microsoft.com/windowsazure/Whitepapers/securityoverview/#02. [Accessed: 15-Sep-2011].

[39] M. Hansen and E. Schlehahn, "TClouds - Privacy meets Innovation," *Symposia*, no. 1/2011, pp. 39-42, 2011.

[40] S. Bugiel, S. Nürnberger, A.-R. Sadeghi, and T. Schneider, "Twin Clouds: Secure Cloud Computing with Low Latency," in *Proceedings of 12th Communications and Multimedia Security Conference*, 2011.

[41] E. Salant, "The RESERVOIR Framework - A Power Tool for CTOs and CIOs," 2011. [Online]. Available: http://62.149.240.97/index.php?page=training. [Accessed: 16-Sep-2011].

10. Appendix A – Exchanged Documents in Secure Trade Lines Project

<u>Exporter (Shipper)</u>

Received: Remittance Advice, Letter of Credit (from Corresponding Bank), Bill of Lading, Dispatch Notice (from Export Forwarder)

Sent: Commercial Invoice (to Corresponding Bank), Commercial Invoice, Packing List (to Export Forwarder)

<u>Corresponding Bank</u>

Received: Commercial Invoice (from Exporter), Bill of Lading (from Export Forwarder), Letter of Credit (from Issuing Bank)

Sent: Remittance Advice, Letter of Credit (to Exporter), Bill of Lading (to Issuing Bank)

<u>Insurance</u>

Received: Insurance Requirements (from Export Forwarder)

Sent: Insurance Contract (to Export Forwarder)

<u>Export Forwarder</u>

Received: Export Approval (from Export Customs), Manifest, Bill of Lading in copy (from Export Agent), Insurance Contract (from Insurance), Commercial Invoice, Packing List (from Exporter), Transport Notify (from Export Transporter)

Sent: Bill of Lading, Dispatch Notice (to Exporter), Bill of Lading (to Corresponding Bank), Export Application (to Export Customs), Bill of Lading in copy, Stowage Plan, Manifest (to Import Forwarder), Instructions, Ship Request (to Export Agent), Arrival Notify (to Export Stevedore), Transport Order (to Export Transporter), Insurance Requirements (to Insurance)

<u>Export Transporter</u>

Received: Transport Order (from Export Forwarder)

Sent: Transport Notify (to Export Forwarder)

<u>Export Customs</u>

Received: Export Application (from Export Forwarder)

Sent: Export Approval (to Export Forwarder)

<u>Export Agent</u>

Received: Manifest (from Export Stevedore), Ship Instructions, Ship Request (from Export Forwarder),

Sent: Load Instruction (to Export Stevedore), Manifest (to Export Port Authority), Bill of Lading in copy, Manifest (to Export Forwarder), Manifest, Ship details (to Export Ocean Carrier), Bill of Lading, Manifest (to Import Agent)

<u>Export Stevedore</u>

Received: Arrival Notify (from Export Forwarder), Load Instruction (from Export Agent)

Sent: Manifest (to Export Agent), Load List (to Export Ocean Carrier)

<u>Export Ocean Carrier</u>

Received: Load List (from Export Stevedore), Manifest, Ship details (from Export Agent)

Sent: Stowage Plan (to Export Port Authority), Manifest, Stowage Plan (to Import Ocean Carrier)

<u>Export Port Authority</u>

Received: Manifest (from Export Agent), Stowage Plan (from Export Ocean Carrier)

Sent:

<u>Import Port Authority</u>

Received: Estimated Time of Arrival (from Import Ocean Carrier), Manifest (from Import Agent)

Sent:

<u>Import Ocean Carrier</u>

Received: Manifest, Stowage Plan (from Export Ocean Carrier), Discharge permit (from Import Customs)

Sent: Estimated Time of Arrival (to Import Port Authority), Estimated Time of Arrival (to Import Agent), Stowage Plan (to Import Stevedore), Cargo Declaration (to Import Customs)

<u>Import Stevedore</u>

Received: Stowage Plan (from Import Ocean Carrier), Manifest (from Import Agent), Transport Order (from Import Transporter)
Sent: Goods Available (to Import Forwarder)

Import Agent
Received: Bill of Lading, Manifest (from Export Agent), Estimated Time of Arrival (from Import Ocean Carrier), Bill of Lading, Clearance (from Import Forwarder)
Sent: Manifest (to Import Port Authority), Manifest (to Import Stevedore), Manifest (to Import Customs), Arrival Notify, Transport Order (to Import Forwarder)

Import Customs
Received: Cargo declaration (from Import Ocean Carrier), Manifest (from Import Agent), Customs Declaration (from Import Forwarder)
Sent: Customs Clearance (to Import Forwarder)

Import Transporter
Received: Transport Order (from Import Forwarder)
Sent: Transport Order (to Import Stevedore), Goods (to Importer)

Import Forwarder
Received: Bill of Lading in copy, Stowage Plan, Manifest (from Export Forwarder), Customs Clearance (from Import Customs), Commercial Invoice, Bill of Lading (from Importer), Goods Available (from Import Stevedore), Arrival Notify, Transport Order (from Import Agent)
Sent: Customs Declaration (to Import Customs), Arrival Notify (to Importer), Transport Order (from Import Transporter), Bill of Lading, Clearance (to Import Agent)

Issuing Bank
Received: Letter of Credit Request (from Importer), Bill of Lading (from Corresponding Bank)
Sent: Bill of Lading (to Importer), Letter of Credit (to Corresponding Bank)

Importer (Consignee)
Received: Bill of Lading (from Issuing Bank), Arrival Notify (from Import Forwarder), Goods (from Import Transporter)
Sent: Letter of Credit Request (to Issuing Bank), Bill of Lading, Commercial Invoice (to Import Forwarder)