SECURING PROCESSES FOR OUTSOURCING INTO THE CLOUD

Sven Wenzel[1], Christian Wessel[1], Thorsten Humberg[2] and Jan Jürjens[1,2]
[1]*Chair of Software Engineering, Technical University Dortmund, Dortmund, Germany*
[2]*Fraunhofer ISST, 44227 Dortmund, Germany*

Keywords: Business Processes, Cloud Computing, Compliance, Risks.

Abstract: Cloud computing is yet one of the leading developments and depicts the biggest progress in web technologies. It offers a convenient way for using shared and easy accessible resources, in both a web-based and demand-oriented sense. However, cloud computing brings concept-based risks, e.g. the risk of private data becoming publicly available. Outsourcing of services into a cloud computing environment arises numerous compliance and security-problems for the potential customer. Legal as well as business requirements have to be met after migration to a cloud environment. Compliance to laws, industry-specific regulations and other rules have to be kept. In this paper we present the research project SecureClouds and our ongoing research towards security and compliance analysis of processes which are to be outsourced into the cloud. We further show a first prototype of an analytic tool-environment that allows us to examine whether outsourcing of a business process is possible while keeping all security and compliance requirements.*

1 MOTIVATION

Cloud computing is currently one of the most rapid growing trends and represents the technological development on the web. Computational power, storage space as well as other complex services are outsourced and made accessible through defined interfaces across the internet. The main advantage is that users are only billed according to the actual usage of the utilized service. Cloud computing hence provides comfortable, demand- and web-based access to shared and freely available resources, which are instantly and automatically made accessible (Mell and Grance, 2009). This enables small- and medium-sized enterprises to minimize costs by configuring their IT infrastructure more efficiently due to a dynamic structure.

At the same time, cloud computing inherently comes along with risks. For example, confidential data could be made public or might be accessed by employees of a cloud service provider. According to different surveys, e.g. (BITKOM, 2009), these risks lead to limited acceptance of cloud computing in business scenarios dealing with confidential data. Especially small and medium-sized enterprises (SMEs) have their doubts regarding cloud computing, although they perfectly fit into the target audience of cloud computing services.

With the research project *SecureClouds* we focus on that target audience and support small and medium-sized enterprises (SMEs) to securely utilize cloud computing technology and thus gaining economic advantages. Our idea is to deal with the risks of cloud computing as early as possible, i.e. in the decision phase whether processes should be outsourced into a cloud environment or not. Therefore, we currently develop a threefold approach to investigate the security and compliance requirements of such business processes. During the decision phase we assist SMEs with a risk analysis for their business processes. The analysis reveals potential security risks for the case that these processes are moved into the cloud. In the implementation phase, we provide compliance and security analysis for the processes, which are partially or even completely executed in the cloud. The analyses are thereby tool assisted. The toolset contains different components: for risk analysis, for compliance analysis and for security analysis. The approach and the toolset are described in what follows. They constitute the ongoing work of

*Parts of this work have been funded financially by BMBF grant 01IS11008D (SecureClouds).

Wenzel S., Wessel C., Humberg T. and Jürjens J..
SECURING PROCESSES FOR OUTSOURCING INTO THE CLOUD.
DOI: 10.5220/0003979306750680
In *Proceedings of the 2nd International Conference on Cloud Computing and Services Science* (CloudSecGov-2012), pages 675-680
ISBN: 978-989-8565-05-1

the BMBF-funded project *SecureClouds*, which also includes the local security consultant *admeritia* and the logistics SME *LinogistiX*. Parts of the approach have already been implemented in prototypes. The validation of the approach is planned to be performed by the logistics service provider LinogistiX, who will serve as test customer for outsourcing business processes into the cloud.

The next section will give a brief overview over *cloud computing* and *compliance*. Subsequently, we present the three pillars of our approach: risk analysis (Section 3), compliance analysis (Section 4), and security analysis (Section 5). The prototype implementation of our approach in a toolset based on the CARiSMA tool environment is shown in Section 6. We briefly discuss related work in Section 7. Finally, we conclude this paper and take a look on open research topics in Section 8.

2 BACKGROUND

Before we introduce our approach on risk, compliance and security analysis of processes that are to be outsourced into a cloud, we first want to introduce some terms.

2.1 Cloud Computing

The American *National Institute of Standards and Technology* (NIST) subdivides cloud computing into three layers named service levels. In a bottom-up manner these layers are *Infrastructure as a Service (IaaS)*, which provides basic virtual hardware resources such as virtual machines or networks, *Platform as a Service (PaaS)*, which acts as a middleware such as a programming interface for distributed and scalable software, and *Software as a Service (SaaS)* representing the top-level layer that provides applications ready to use.

Depending on the chosen layer type, different security needs have to be considered. If an IaaS structure was chosen, only the bare virtual hardware is provided. Therefore all protection needed must be installed by the customer of the cloud service. With PaaS it is necessary that the software developed on top of it meets the security standards needed for the business process. On the SaaS layer software must be chosen, that comply with the security needs.

Furthermore a cloud belongs to one of four characteristics defined as the deployment model, meaning its accessibility or the intended user groups of the cloud: *Private clouds* are only accessible for the users of one specific company or organization. *Community clouds* are shared between several companies or organizations which have common goals or security requirements. *Public clouds* are public accessible for a large range of users. Typically a cloud provider chooses this deployment model to sell cloud services. *Hybrid clouds* are composed of two or more clouds of different deployment models. This model may be used for load balancing e.g. if the resources of a private cloud are insufficient in times with high peaks, a public cloud can be hired for compensation.

Depending on the deployment model different security needs have to be considered. While private clouds can be considered relatively secured against attacks from outside they are exposed to internal attackers. The security needs for community clouds are a bit more complex, since there is more than one party accessing the cloud. The weakest security level is found within public clouds. Therefore a business process outsourced to a public cloud must either be non-critical or they deserve respective data protection. Carrying sensitive data to or from a private cloud to a public or community cloud as it happens in hybrid clouds, has also to be realized in a secured manner.

2.2 Risks & Compliance

In our understanding risks are defined as the components mentioned in the *BSI basic protection catalog* (BSI, 2006); probability of occurrence and amount of damage are seen as irrelevant. Risks are considered to be more IT security related while compliance is defined as the internal regulations and laws a company have to comply with. Security risks, compliance and security are tightly bound to each other, e.g. consider an attacker that is able to gather personal data from an unsecured server. Because of the security issue the compliance requirement of personal data protection cannot be held.

Risks. Here, a risk is defined as a IT security related weakness in a cloud-based process. This may be an unsecured communication channel between one host of the cloud with another, insufficient rights management on a specific host, or just the processing of confidential data.

Compliance. A security analysis on the cloud computing environment should be performed before a security requirement analysis is performed on the business operation. This will yield the maximum number of security requirements that can be met. The conformance to compliance regulations should be audited on three levels.

Process & Compliance Analysis. Documents from which business processes can be derived should be

analyzed. Our approach, which is presented in what follows, considers processes given in form of process models (e.g. UML activity diagrams or BPMN models). The risk analysis works also on less structured documents such as textual process descriptions.

Design Time Compliance. The implementation of a business process has to comply with legal regulations and company policies. The main focus here is the cloud interface, the process steps within the cloud and the data flows between the cloud and the user.

Runtime Compliance. It has to be ensured that all compliance-relevant and critical processes (esp. those outsourced into a cloud) are monitored and logged. Such a monitoring can be performed by using business process mining and conformance checking (W. van der Aalst et al., 2007). It is not further considered in this paper.

3 RISK ANALYSIS

The first pillar of our approach is a security risk analysis of the business processes that are to be outsourced into a cloud environment.

We have developed an algorithm that enables enterprises to check their business processes for activities that comprise security requirements. The algorithm takes a business process as input. The entire text of the model is extracted from all entities of the model and is then imported in an internal data model.

The other input for the algorithm is a set of *security patterns*, which are derived from the basic protection catalogs of the BSI (BSI, 2006). A security pattern consist of three parts:

1. An unique *Identifier*, derived from the hierarchy given by the catalogs.
2. The *Title* stating the content of the pattern.
3. The complete text itself, as given in the standard.

Using methods from the field of natural language processing, the algorithm is able to compare the descriptions of an activity to security patterns, thus identifying relevant activities and patterns. Given a pair of one activity and one pattern, the method proceeds in three steps (cf. Algorithm 1):

First, all *stop words* are removed from the description text. Those words are a concept from the area of language processing and form a set of words that do not contribute to the meaning of a text. Typical examples are words like ”the”, ”that” or ”a” (Runeson et al., 2007). As those are quite common in the texts of the patterns as well as in the activity labels, these would cause many irrelevant matches and are therefore excluded from further consideration.

Algorithm 1: Risk detection algorithm.

Input: Text T, composed of words $W = \{w_1, \ldots, w_n\}$, Pattern P=$\{p_1, ..., p_n\}$, Set S of stopwords
Result: Set M of security-relevant words in T

$W = \bigcup_{w \in W} w \cup synonyms(w)$ /* extend */
$P^- = P/S$ /* patterns w/o stopwords */
$M = W_{ext} \cap P^-$ /* matching words */
return M

The next step applies language tools to the description: For each word remaining after the removal of the stop words, words with similar meanings are added, to allow for a better comparison to the patterns.

Finally, the set of words of the description after the synonyms were added is compared to the set of words of the pattern. Those words that are elements of both sets are detected, if the number of these words exceeds a fixed threshold, the considered activity is classified as security relevant. The classified activities are currently listed in a report. Visualization such as coloring the tasks in the BPMN diagram according to the severity would be possible.

4 COMPLIANCE ANALYSIS

Once the decision is made and the business process (or parts of it) is outsourced into the cloud, the process often requires some adaption to the new environment. I.e. tasks are distributed to different actors or ordering is rearranged for a more smoothly execution.

The second pillar of our approach is a compliance check that verifies whether the legal regulations (e.g. binding of duty) are still kept. Therefore, we currently develop a catalog of compliance rules that can be checked on a business process model.

Simple rules can easily be expressed as OCL constraints. An example is the *binding of duty* constraint, which ensures that two tasks are performed by the same actor (or entity):

context TaskSet **inv:** self.selectedTasks
$\rightarrow$ forAll($x, y | x <> y$ implies x.performer
$\rightarrow$ forAll($z | y$.performer $\rightarrow$ includes(z)))

In order to formulate such constraints, the meta model of the business process model requires some extensions. In case of the above-mentioned binding of duty, we must be able to define the set of tasks that are bound (i.e. TaskSet). We have developed an extension model for BPMN. It allows to define additional information such as role bindings, tasks set, etc. and can be weaved into an BPMN model.

More complex rules, especially those containing

flow characteristics such as 'Task A has to be performed before Task B', can hardly be expressed in OCL. Therefore we propose to define process fragments and rule sets that can be compared to the process. A fragment is basically a snippet of a process necessary to express a certain rule. The compliance rules can be checked by analyzing whether the structure of a given process fragment is contained in the business process, similar to a graph isomorphism check. However, it is not a real isomorphism since the business processes may contain additional tasks and the same process can be modeled in different ways.

For the first prototype of the compliance analysis, we have implemented the *Minimum Requirements for Risk Management for Insurance Undertakings (MaRisk VA)* published by the German Federal Financial Supervisory Authority (BaFin) with regards to the *Solvency II* ordinance. They define for instance that the executive board of an enterprise has to be informed if significant problems are discovered during an audit. We systematically collected the process requirements defined in MaRisk and formulated nine process fragments and 40 rules out of that.

Each check analyzes the existence of the significant tasks of a process fragment in the business process that is analyzed. The rules inspect certain properties such as existence of paths between certain tasks, ordering of tasks, or boundary events. The mapping between tasks of the process fragment and tasks of the analyzed process is still a manual task. However, the manual effort is reduced to locating certain tasks, the analysis of the relationships and interactions can be automatized.

5 SECURITY ANALYSIS

The third pillar of our approach is the automated analysis of security properties. The goal is to integrate process models and UML deployment diagrams within the analysis process. This enables the security analysis of the (physical) distribution of a process and of the communication between different entities.

Deployment diagrams can be used to model the cloud environment, i.e. the different virtual (and real) machines, the provided services, and so on. Therefore we currently develop a UML profile that provides stereotypes to classify the different components, e.g. `<<IaaS>>`, `<<PaaS>>`, or `<<public cloud>>`. Tagged values can be used to assign additional information to the entities, such as geographical location of a server. By combining the business process model and the deployment model we are able to determine which parts of the analyzed business process will be

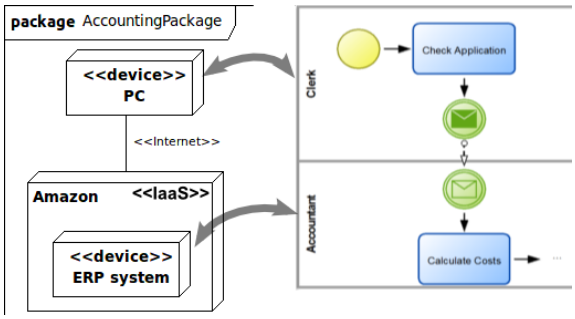

Figure 1: Distribution of processes onto system nodes.

deployed into the cloud and how (i.e. which service levels are used etc.). Therefore we need to map the elements of the process model onto the elements of the deployment model. The mapping can be performed by the system designer. An example is illustrated in Figure 1. The mapping allows us to classify parts of the business processes into groups with different security needs, e.g. tasks processing confidential data that are delegated to a service provider. According to the security need, the different parts of the processes can be proposed to the user for further inspection.

Besides that, it enables the automatic inspection in form of model-based security analysis. An example is the data protection prescribed by the German data protection law, *Bundesdatenschutzgesetz (BDSG)*, §4b. It says that personal data must not be stored or processed outside the European Union. Hence, if a task processes personal data such as the social security number, we have to ensure that the machine which hosts this tasks is located within the EU. Given the integrated process and deployment model such properties can be checked easily, because the mapping between processes and deployments shows us, which tasks are executed on which system node.

Another aspect of our approach is to inspect the communication between the parts of the business process. Especially the communication between parts hosted in the cloud and the other are of interest. This information can also be extracted from the deployment diagram. Using UML deployment diagrams has the advantage that it can be annotated with *UMLsec* (Jürjens, 2005) stereotypes to emphasize security related connections. UMLsec is an extension for UML that allows us to annotate models with security related properties, e.g. a communication link between two computers can be marked with `<<encrypted>>`, to denote that the link if established has to be encrypted. Another possible idea is to annotate the business process model with security related information. This way security checks can be accomplished directly on the business process level and no mapping would be necessary as a preliminary step.

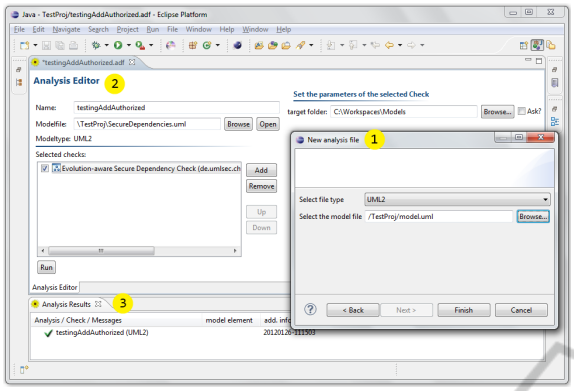

Figure 2: The toolset user interface.

6 TOOL SUPPORT

We have implemented the risk analysis and the compliance checks as plugins for the CARiSMA framework[2]. It is fully integrated into the Eclipse GUI (see Figure 2). CARiSMA provides an analysis wizard (1) for the definition of a model analysis. The user can select the model(s) he wants to analyze and adjust parameters of the different checks (e.g. which ontologies to use). The analysis editor (2) allows the user to modify the settings of existing analyses. The results of an executed analysis are displayed in the analysis results view (3). Different icons indicate whether a model contains risks or compliance violations.

The back-end of CARiSMA is build on top of the Eclipse Modeling Framework, which implements the OMG Meta Object Facility (MOF) specification. For UML models, the Eclipse UML2 implementation and the UMLsec profile (Jürjens, 2005) is used. For BPMN models we use the Eclipse BPMN project which we have extended by additional classes for inserting security related information such as role assignments (Michel, 2011).

The risk analysis and the compliance analysis have been implemented as different plugins (i.e. checks) for CARiSMA. The security analysis will be implemented in the same manner. The risk analysis contains of four major components: An *extractor* extracts the vocabulary of a process model. A *normalizer* cleans the extracted information and eliminates the different flections of words. An *expander* extends the normalized information, e.g. with synonyms, based on powerful ontologies. The *analyzer* searches for the different security and risk patterns in the normalized and extended texts. It can use different pattern repositories we have implemented such as the BSI basic protection catalog (BSI, 2006) or ISO 27000. Different normalizer implementations can extract from different sources and decouple our approach from certain modeling languages. For each language a certain normalizer can be implemented. Different expanders allow us to include certain ontologies to find synonyms etc. The current implementation uses standard language libraries such as *Wordnet*[3]. Custom security- and cloud-related ontologies are currently developed.

The compliance analysis is basically realized as a set of checks, each implementing a certain rule. A general check using the OCL interpreter of the Eclipse project, even enables the quick definition of new constraints. So far UML activity diagrams and BPMN models are supported. The compliance analysis based on process fragments (i.e. the MaRisk rules from Section 4) is currently limited to BPMN models.

The security analysis exists so far only on conceptual level. However, since CARiSMA is the successor of the UMLsec tool, a broad range of UML-based security checks is available. The security analysis for UMLsec models can be easily adapted to the cloud-specific analysis proposed in Section 5, especially because basic functionality such as the `<<secure links>>` check are already given.

7 RELATED WORK

Cloud computing is still on the peak of Gartner's technology hype cycle (Dixon and Jones, 2011). Especially the economic potential of cloud computing for small and medium-sized enterprises has been discussed for quite a while now (BITKOM, 2009). However, if it comes to confidential information such as enterprise data and other security-related issues, the majority of SMEs is still doubting.

A project related to our approach and the project *SecureClouds* is *CloudCycle*[4]. It focuses on cloud providers and offers services that allow them to guarantee their customers that they are compliant with security policies and further regulations. The approach of CloudCycle is a suitable complement for our approach. Once business processes are successful outsourced into the cloud their security and compliance can be monitored.

Ontologies for cloud computing and cloud security have been presented by (Gräuler et al., 2011). They analyzed the different sources of risks within cloud computing environments and manifested them in an ontology. Based on that ontology, they provide a database of cloud providers that allows users

[2]http://carisma.umlsec.de

[3]http://wordnet.princeton.edu/

[4]http://www.cloudcycle.org

to select a providers based on certain security properties. This is especially interesting for finding a suitable cloud provider after potential risks of a business process have been revealed by our approach.

Further work on general IT risk analysis exists. (Peschke et al., 2011) present the *RiskFinder* which is a precursor of our risk analysis component. It analyses UML models with respect to security relevant vocabulary. Schneider et. al. propose a heuristic search based on Bayesian filters (Schneider et al., 2011). HeRA realizes a feedback-driven approach for security analysis during requirements engineering (Knauss et al., 2009). These approaches provide powerful rules, however, they work only on single words and do not consider language databases.

An approach to encode and check security requirements in BPMN models has been presented in (Wolter et al., 2008). However, these requirement focus only on closed systems and are not eligible for open processes which are meant to be executed in cloud environments. Security requirements on service orchestration level have been discussed in (Menzel et al., 2009).

8 CONCLUSIONS & OUTLOOK

In this paper we have presented our research project *SecureClouds* which develops an approach to assist small and medium-sized enterprises for deciding which of their business processes are eligible to be outsourced into a cloud computing environment. The approach is based on three pillars. Firstly, risk analysis is used to unfold potential risks of business processes that are to be outsourced into a cloud environment. Secondly, compliance analysis allows the enterprises to check whether the processes are still compliant after adapting them to cloud environments. Finally, a security analysis enables the validation of security properties of cloud-based business processes. The approach is currently being implemented in a toolset based on the CARiSMA analysis tool environment which is a framework to provide a broad collection of different model-based security analyses.

While the approach presented here only focuses on the perspective of the users of cloud environments, it would be interesting to inspect also the business processes within the cloud provider's domain. Furthermore, the analysis of cloud environment themselves might be interesting for inspecting the influence of cloud architectures on security properties.

The approach presented here is still in an early stage of development. It is the result of the first year of the project *SecureClouds*. For the second year of that project, we plan to deepen our research in different ways. One major aspect will be the further development of the ontologies that we use for risk analysis. Additional compliance and security checks are also planned to be implemented. Last but not least, the most important step we achieve is the evaluation of our approach in real case study in the logistics domain together with the enterprise partners of the project.

REFERENCES

BITKOM (2009). Cloud-Computing - Evolution in der Technik. Technical report, BITKOM.

BSI (2006). IT Basic Protection Catalog. *Online: http://www.bsi.bund.de*.

Dixon, J. and Jones, T. (2011). Hype cycle for business process management. Technical report, Gartner Study.

Gräuler, M., Martens, and B.; Teuteberg, F. (2011). IT-Sicherheitsmanagement im Cloud Computing. In *Proceedings INFORMATIK 2011*, Germany.

Jürjens, J. (2005). *Secure Systems Development with UML*. Springer, 1. edition.

Jürjens, J. and Shabalin, P. (2007). Tools for secure systems development with UML. In *International Journal on Software Tools for Technology Transfer (STTT), Volume 9 (5-6): 527-544*.

Knauss, E., Lubke, D., and Meyer, S. (2009). *Feedback-driven requirements engineering: The Heuristic Requirements Assistant*. In *ICSE'09*, Washington, DC.

Mell, P. and Grance, T. (2009). Effectively and Securely Using the Cloud Computing Paradigm.

Menzel, M., Thomas, I., and Meinel, C. (2009). Security requirements specification in service-oriented business process management. In *ARES*.

Michel, M. (2011). Konzeption und Umsetzung eines UMLsecTool-Plugins zur Prüfung von Authorization Constraints für die Prozessmodellierungssprache BPMN 2.0. Bachelor thesis, TU Dortmund, Germany.

Peschke, M., Hirsch, M., Jürjens, J., and Braun, S. (2011). *Werkzeuggestützte Identifikation von IT-Sicherheitsrisiken*.

Runeson, P., Alexandersson, M., and Nyholm, O. (2007). Detection of duplicate defect reports using natural language processing. In *ICSE'07*, Washington, DC.

Schneider, K., Knauss, E., Houmb, S., Islam, S., and Jürjens, J. (2011). Enhancing security requirements engineering by organizational learning. *Requirements Engineering*, pages 1–22.

W. van der Aalst, H. Reijers, A. Weijters, F. van Dongen, M. Song, H. Verbeck. (2007). Business process mining: An industrial application. *Information Systems, Vol. 32, No. 5.*

Wolter, C., Menzel, M., and Meinel, C. (2008). Modelling security goals in business processes. In *Modellierung 2008*, Germany.