ARCHITECTURE FOR COMPLIANCE ANALYSIS OF DISTRIBUTED SERVICE BASED SYSTEMS

Jonathan Sinclair, Benoit Hudzia, Maik Lindner
SAP Research, Belfast, Northern Ireland, U.K.

Alan Stewart, Terry Harmer
School of EEECS, Queen's University Belfast, Belfast, Northern Ireland, U.K.

Keywords: Compliance, Auditing, Enterprise cloud computing, Data protection.

Abstract: Businesses today are required to comply with a litany of legislation, regulations and standards. However, with an increasing utilisation of the internet for delivering products as services, challenges arise in assessing and maintaining compliance. We propose to define an architecture that attempts to leverage the dynamism of service-based infrastructures in order to process the real-time compliance state of a system.

1 INTRODUCTION

With the advancement of web-based infrastructures it is perceived that computing resource will become the 5th utility after water, electricity, gas and telephony (Baumann et al., 2010; Buyya et al., 2008). Business economics is the main driver of this transformation which enables companies to radically transform their business processes and operations. It allows them to streamline the delivery and consumption of their products and solutions over a network infrastructure. This new business model and mechanisms for integrating services and IT resources in a seamless and ubiquitous is known as Internet of Services (IoS) (Heuser et al., 2008). IoS endeavours to reduce TCO for customers by making complex software a commodity (Janiesch et al., 2009). IoS provides a business model in which IT is evolving into a service driven ecosystem with outsourced distributed computing resources, such as cloud, providing an economical way of acquiring hardware.

However, companies willing to leverage this new business model have to abide by the current state of legislation which hampers its adoption even though cloud offers benefits such as elasticity and rapid deployment, improving companies' efficiencies in times of economic hardship. The risk and financial penalty associated with non-compliance is too great for businesses to ignore. The main source of these legal issues is third-party trust, as companies are ultimately responsible for demonstrating data compliance.

This research contributes to reducing the legal-technical issues that hinder adoption of cloud-based environments for enterprises, by addressing the assurance of legal risk by way of facilitating compliance auditing. In this paper we propose an architecture for enabling an automated semi real time monitoring, audit and compliance service (MACS) for verification of services provided within cloud environment.

The MACS is provided with the specification of the compliance associated with the services and is capable of observing and logging the relevant service generated events, in order to determine if the actions, events are consistent within the domain of compliance associated. MACS rely on a set of rules that provides constructs to specify what rights, obligation and prohibitions become active and inactive after the occurrence of events related to the service lifecycle. Our solution is specifically targeted for audit and compliance verification of composite services within cloud environments.

We will first discuss in the subsequent section the fundamentals of auditing and compliance. Then we will describe a simple and composite service use case and the difficulty associated them . In Section 4, we will identify and generalize the challenges of auditing the compliance of regulations for distributed service based architectures. Finally, we outline the design for an auditing service that leverages the capabilities of such architecture.

Sinclair J., Hudzia B., Lindner M., Stewart A. and Harmer T..
ARCHITECTURE FOR COMPLIANCE ANALYSIS OF DISTRIBUTED SERVICE BASED SYSTEMS.
DOI: 10.5220/0003448702860292
In *Proceedings of the 1st International Conference on Cloud Computing and Services Science* (CLOSER-2011), pages 286-292
ISBN: 978-989-8425-52-2

2 FUNDAMENTALS OF AUDITING AND COMPLIANCE

In this section we will first introduce the fundamental of auditing then provide compliance background.

2.1 Background of Auditing

Auditing of information systems is the process of collecting and evaluating evidence to determine whether it safeguards assets, maintains its data integrity while achieving organizational goals effectively and consumes resources efficiently (Weber, 1998). Traditionally auditing is a semi-manual operation where the data is collected for a determined time period and then analyzed according the legal / regulatory requirements in order to determine whether or not statutory requirements have been met. This process is time-consuming and error-prone, and for most companies, compliance audits are carried out on static and rigid services. However, as the IT service move to a more dynamic and distributed environment the classic auditing process cannot be applied anymore.

2.2 Background of Compliance

Compliance is defined as being in accordance with relevant government or industrial legislation, regulations, and standards. Companies breaching compliance rules incur either or both legal and financial penalties. Due to this risk, compliance management has become very important and technology is having a ever increasing significance and impact on virtually every phase of the audit process (Flint, 2009; Janvrin, 2007). As a result it becomes critical with the increased reliance on cloud computing for organizations and consumers to be aware of their responsibility related to determining who is accessing data, what actions are being performed and where data is stored (Pearson, 2009; Moreau et al., 2008; Morrison et al., 2000).

The complexity of such compliance requirements is increased due to the platform and infrastructure details being abstracted making them invisible to the service model (SaaS). As a result it becomes extremely difficult to take full responsibility for who can access data, who sees it and how it is stored, since the premise of the cloud is that customers don't necessarily need to know or care where their data is (Wood, 2009). This paradigm being in contradiction with compliance legislation such as the EU Directive (Parliament, 1995) which require companies to be aware of the jurisdiction in which their data is processed. As a result the necessity to provide third party compliance verification system emerges in order to remove the burden and facilitate the adoption of a service consumption model.

2.3 Related Work

Current approaches to auditing involve the aspect of managing governance, risk and compliance (Silveira et al., 2010; Cederquist et al., 2007). Typical auditing methodologies require the most relevant data related to an event. However this approach does not ensure that the legislation/regulation has been upheld. In order to be able to conduct an audit of system compliance it is necessary to have a machine interpretable formalisation of the legislation, previous research has been done in this area (Baumann et al., 2010; Conrad et al., 2007). Typical auditing tools, however, rely on the manual input of data. In contrast by applying software such as complex event processing (CEP), events can be filtered, transformed and aggregated to provide new interpretations of data that were not previously possible (Etzion and Niblett, 2010). These formalisations can be used in conjunction with Service Level Agreements (SLA's) to provide input for the compliance auditing architecture (Brandic et al., 2010; Skene, 2007).

3 CLOUD COMPLIANCE CHALLENGES

As the customer base only consumes the final product over internet connection, the geographical locality of services provided through cloud is relatively limited. As a result, the cloud provider may choose the geographic placement of data centres based on various cost benefits, including energy. However the physical location of data being accessed, stored, processed or transferred is of critical importance to the applications of data protection legislation such as EU Directive (Parliament, 1995). Hence, cross-country legal aspects become common place and are a major challenge for IoS. Data transfer related legislation may or may not be enforced or reported depending on the source or destination of the transfer (Jaeger et al., 2009). Geographic locality challenges audit and compliance in the following forms:

3.1 Cross-jurisdictional Services Enforcement

Multi-national companies experience increasing difficulty simultaneously complying with a number of conflicting data protection requirements at the same time. Enforcing geographical deployment of a service within a jurisdiction that meets all requirements requires an efficient compliance evaluation system for the assessment of distributed services.

By example, a company may deploy a composite CRM system. This system is split and the database is stored separately to comply with financial compliance requirement for processing and storing data on independent physical machines within the same geographical region (Council, 2004; 107th United States Congress, 2002; 106th United States Congress, 1999). However, the company is also required to backup this data in a different geographic region (Swanson et al., 2010). As a result, the company needs to guarantee the physical distribution of its service and associated backup.

3.2 Performance and Availability Enforcement

Services may have to be deployed in data-centres of a particular geographical region in order to satisfy legal conditions. The resulting deployment might restrict the performance / availability between the deployed service and the user which can result in breach of defined SLAs.

3.3 Disaster Recovery/Backup Implications

Legislation (Law, 2000) and regulatory requirements for various industries (Swanson et al., 2010; on Banking Supervision, 2009) define that the data be suitability backed up for disaster recovery purposes in a different geographic region. On the other hand, other legal conditions can restrict deployment to a solitary geographic region, therefore a major confliction is imposed, in which the consumer is required to take remedial action.

3.4 Data Accessibility Aspects

Data access is another point of contention with respect to compliance. Who can access data? What data can be accessed? How should data be accessed? In IoS the aspect of service auditing must consider the compliance requirements of all consumers in terms of both company and systems multi-tenancy (Alliance, 2009; Mell and Grance, 2009; Sotto et al., 2010).

User Multi-tenancy. A company may accept to deploy their system virtually co-located on the same physical machine as other companies' deployments (Li et al., 2010). This may raise concern in the context that data from both companies is stored on the same physical device, regulations exist forbidding such setup for critical customer or company data.

Service and Systems Multi-tenancy. To achieve efficiency and cost-savings, a company can co-locate several virtual systems on the same physical resource. However certain regulations such as the Payment Card Industry Data Security Standard (PCI-DSS) (Council, 2004), specifies under requirement 3 that cardholder data needs to be stored and processed independently in different physical devices. Therefore as defined in the use case a composite on-demand service would be best suited in fulfilling these requirements. However this scenario has many aspects of compliance that require to be ensured. By example, a cloud provider may migrate a virtual machine hosting a service to a physical server already hosting another service that should not be co-located with one another. As a result, such action undertaken without the customer awareness will breach the compliance of both services.

4 USE CASE - CUSTOMER RELATIONSHIP MANAGEMENT

In this section we highlight the compliance aspect a business is required to make when deciding on the deployment method of a service. There are 3 deployment scenarios, on premise, on demand and hybrid. On top of the deployment model, a service can be either a single atomic service or a composition of smaller services exposed as a unique one. Table **??** shows the various compliance areas that need to be considered for each variation of the use case we describe in the following paragraph.

We use a typical CRM application as a foundational base for our scenario. We first analyse the interactions in order to determine their relevance for compliance monitoring for the 3 following deployment scenarios Figure 1:

1. The CRM and respective database is deployed on premise. As a result, both the customer and employee interact directly with the system. All data access, processing, storage and transfer are managed within the system which makes any physical access to the system a controlled variable by the company.

2. For hybrid deployment, the CRM and database are deployed independently, one on premise and the other as a remote service. In this scenario not only does the compliance data access, processing and physical access have to be managed in two distributed sites but also compliance of both data transfer between sites and the geographical locality of the remote service are important.
3. Finally with full on-demand procurement there are two scenarios to consider, deployment of the CRM and database as a singleton, or independently. In both scenarios compliance of data access, processing, storage and physical access are required, but data transfer is only important in the case of service composition.

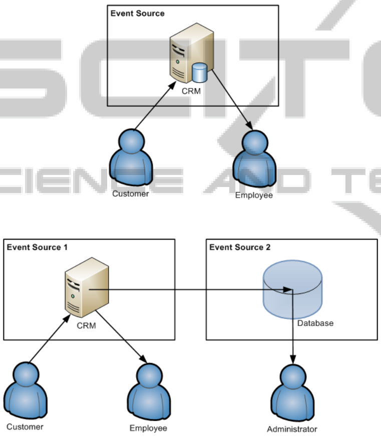

Figure 1: Use Case Deployment.

Service composition is seen as a key benefit for IoS. By composing a selection of simple services to create a more complex offering, companies can create and offer specialized services dynamically, adding new economic value (Blau et al., 2009). This flexibility, however, creates new legal challenges for auditing compliance, as this new service has to adhere to a composition of compliance requirements. In certain extreme cases, two compliance regulations can have contradictory interpretations. Therefore at the service composition design stage an agreement for the compliance requirements of the new service has to be derived in order to prevent conflict in the auditing process. Moreover, composite services inherently reduce the visibility within the internal operation of the composite. As a result the audit process suffers from an increased complexity due to the difficulty of tracing and verifying the origin of the information. By example, in the use case we presented: the CRM system stores the personal data collected by the company. As a result, the consumer may not be able to control who can access this information.

5 COMPLIANCE AUDITING ARCHITECTURE

Figure 2 represents the architecture of MACS. MACS architecture is comprised of five distinct layers, each managing a different aspect of the audit and compliance engine:

Event Source. The data and logs returned from various logical, physical or virtual components in the system. The source may be a sensor, application, messaging framework, business process, data store, client applications. Each source is authenticated and uses secure means of communication. In the best case scenario every source would be verified using a system similar to trusted platform computing (TPC) (Santos et al., 2009).

Event Transport. Typically an enterprise service bus (ESB) (Schmidt et al., 2005) controls how data is routed to the event processing engine, in a standardized format. This allows the MACS system to have a reliable and uniform delivery system that can be used as an interface between the source and the MACS but also for internal communication.

Event Processing Engine. This layer processes events in three levels each with increasing processing complexity but decreasing event throughput. Each level is a filter and information enhancer for the next one. By removing unwanted events, aggregating them we increase the degree of information and lower the amount of data to be processed.

- Anomalous Filtering. Removes data that is not relevant to the compliance process. The operations at this level receive a high throughput of data or event rate but are of low complexity and will be typically executed by complex event processing engine (Mulo et al., 2010; Rozsnyai et al., 2007).
- Temporal Filtering. Synchronizes time and event type inconsistencies and correlates events, aggregating data over a window of time. The operations at this level receive medium throughput of data or event rate but are of medium complexity.
- Compliance Filtering. Event streams are compared and evaluated against business rules that have been derived from the legalization for which compliance is being assessed. The operations at this level receive low throughput of data or event rate but are of high complexity. This level will be using rule based engine (Chesani et al., 2009) as well as more complex solution as case based

Table 1.

Service Type	On Premise	Hybrid	On Demand
Singleton	Physical Access, Data Access and Retention	N/A	Geographic Locality, Physical Access, Data Access and Retention
Composite	Physical Access, Data Access and Retention	Geographic Locality (remote service only), Physical Access, Data Access, Data Transfer and Retention	Geographic Locality ,Physical Access, Data Access, Data Transfer and Retention

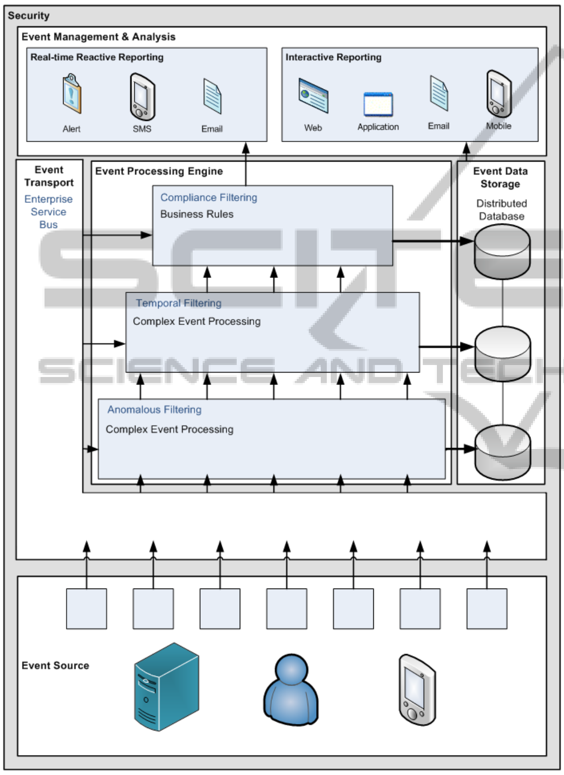

Figure 2: Logical Architecture.

evaluation engine and semantics evaluation tools - (Elgammal et al., 2010).

However as we increase the value of the information as we move up the layer the amount of processing increases as we decrease the throughput of information to process. The advantage of this approach means that we are able to efficiently handle the data storm of events and logs pushed while intelligently filtering and cherry picking only the relevant information for (semi) real time auditing and compliance checks.

Event Storage. MACS stores all event data in persistent storage such as Hadoop and HBase (Zhang et al., 2010) for storage, log analysis, and pattern discovery/analysis. This enables us to process historic queries and results from event correlation which enable us in future works to provide predictive analysis for early warning of compliance deviation.

Event Management and Analysis. Queries can be defined by the user and can be compared in run-time or on historic data. Note as mentioned in section 2.3. We do not aim to automatically translate compliance law and contract in rules but aim to leverage previous work done in the field in order to apply it to our system. Monitoring features set by users enable them to trigger alerts if compliance is not met. These alerts can be sent by either SMS, email or mobile application. Finally more traditional audit reports can be produced as well as historic data or predictive trends. The event source, processing and storage layers in this architecture is implemented in a distributed approach, in order to tackle the challenges highlighted in terms of geographic locality. By example, a portion of the architecture can be deployed on premise in order guarantee the confidentiality of the data while exposing only the necessary information to external compliance engine.

5.1 Architecture Benefits

The CEP-based architecture offers major advantages compared with traditional monitoring techniques. The event-driven approach provides flexibility in it's loosely coupled design in which both the integration of multiple event sources and the filtering and standardisation of different event types can be easily addressed. Integration of event sources is also eased by use of a common event transport (enterprise service bus).

5.2 Architecture Challenges

The management of event publication, subscription and filtering may be difficult. Identifying the relevance of the data required and retrieving it, in order to determine the outcome of legal requirements will be key to the accuracy of processing business rules.

6 CONCLUSIONS

We have highlighted how businesses are under increasing pressure to manage the litany of legislation and regulation. Coupled with the adoption of web-based infrastructures and composite services. We propose an architecture that leverages the dynamism of service-based infrastructures and enables real-time compliance processing. The described architecture is loosely based on an event-driven service-oriented architecture (SOA). The loose coupling allows for easy scalability and distribution of both event processing and storage components whilst managing processing complexity. This forms a foundation for further research into a compliance-driven auditing architecture for distributed systems.

REFERENCES

106th United States Congress (1999). Gramm-leach-bliley act.

107th United States Congress (2002). Sarbanes-oxley act. Securities and Exchange Commission.

Alliance, C. S. (2009). Security guidance for critical areas of focus in cloud computing. http://www.cloudsecurityalliance.org/csaguide.pdf.

Baumann, C., Peitz, P., Raabe, O., and Wacker, R. (2010). Compliance for service based systems through formalization of law. In Filipe, J. and Cordeiro, J., editors, *Proceedings of the 6th International Conference on Web Information Systems and Technology*, volume 2, pages 367–371, Valencia, Spain. INSTICC Press.

Blau, B., Kramer, J., Conte, T., and Dinther, C. v. (2009). Service value networks. In *Proceedings of the 2009 IEEE Conference on Commerce and Enterprise Computing*, pages 194–201, Washington, DC, USA. IEEE Computer Society.

Brandic, I., Dustdar, S., Anstett, T., Schumm, D., Leymann, F., and Konrad, R. (2010). Compliant cloud computing (c3): Architecture and language support for user-driven compliance management in clouds. *Cloud Computing, IEEE International Conference on*, 0:244–251.

Buyya, R., Yeo, C. S., and Venugopal, S. (2008). Market-oriented cloud computing: Vision, hype, and reality for delivering it services as computing utilities. In *HPCC '08: Proceedings of the 2008 10th IEEE International Conference on High Performance Computing and Communications*, pages 5–13, Washington, DC, USA. IEEE Computer Society.

Cederquist, J., Corin, R., Dekker, M., Etalle, S., den Hartog, J., and Lenzini, G. (2007). Audit-based compliance control. *International Journal of Information Security*, 6:133–151. 10.1007/s10207-007-0017-y.

Chesani, F., Mello, P., Montali, M., Riguzzi, F., Sebastianis, M., and Storari, S. (2009). Checking compliance of execution traces to business rules. In Aalst, W., Mylopoulos, J., Sadeh, N. M., Shaw, M. J., Szyperski, C., Ardagna, D., Mecella, M., and Yang, J., editors, *Business Process Management Workshops*, volume 17 of *Lecture Notes in Business Information Processing*, pages 134–145. Springer Berlin Heidelberg.

Conrad, M., Funk, C., Raabe, O., and Waldhorst, O. (2007). A lawful framework for distributed electronic markets. In Camarinha-Matos, L., Afsarmanesh, H., Novais, P., and Analide, C., editors, *Establishing The Foundation Of Collaborative Networks*, IFIP International Federation for Information Processing, pages 233–240. Springer Boston.

Council, P. C. I. S. S. (2004). Payment card industry data security standard.

Elgammal, A., Turetken, O., Heuvel, W. v. d., and Papazoglou, M. (2010). On the formal specification of business contracts and regulatory compliance. Open access publications from tilburg university, Tilburg University.

Etzion, O. and Niblett, P. (2010). *Event Processing in Action*. Manning Publications.

Flint, D. (2009). Law shaping technology: Technology shaping the law. *International Review of Law, Computers & Technology*, 23 , 1:5–11.

Heuser, L., Alsdorf, C., and Woods, D. (2008). *International Research Forum 2007*. Evolved Technologist Press.

Jaeger, P., Lin, J., Grimes, J., and Simmons, S. (2009). Where is the cloud? geography, economics, environment, and jurisdiction in cloud computing. *First Monday*, 14:5.

Janiesch, C., Niemann, M., and Repp, N. (2009). Towards a service governance framework for the internet of services. In *17th European conference on information systems (ECIS)*, pages 1 –13.

Janvrin, D. (2007). The impact of information technology on the audit process: An assessment of the state of the art and implications for the future. *Managerial Auditing Journal*, 16:159–164.

Law, U. S. P. (2000). Health insurance portability and accountability act.

Li, X.-Y., Shi, Y., Guo, Y., and Ma, W. (2010). Multi-tenancy based access control in cloud. In *Computational Intelligence and Software Engineering (CiSE), 2010 International Conference on*, pages 1 –4.

Mell, P. and Grance, T. (2009). Effectively and securely using the cloud computing paradigm. National Institute of Standards and Technology.

Moreau, L., Groth, P., Miles, S., Vazquez-Salceda, J., Ibbotson, J., Jiang, S., Munroe, S., Rana, O., Schreiber, A., Tan, V., and Varga, L. (2008). The provenance of electronic data. *Commun. ACM*, 51, 4(4):52–58.

Morrison, R., Balasubramaniam, D., Greenwood, M., Kirby, G., Mayes, K., Munro, D., and Warboys, B.

(2000). An approach to compliance in software architectures. *Computing and Control Engineering Journal*, 4:195–200.

Mulo, E., Zdun, U., and Dustdar, S. (2010). Monitoring web service event trails for business compliance.

on Banking Supervision, B. C. (2009). *International Convergence of Capital Measurement and Capital Standards*. Bank for International Settlements Press & Communications CH-4002 Basel, Switzerland.

Parliament, E. (1995). Directive 95/46/ec of the european parliament and of the council. Official Journal of the European Communities. http://ec.europa.eu/justice/policies/privacy/docs/95-46-ce/dir1995-46_part1_en.pdf.

Pearson, S. (2009). Taking account of privacy when designing cloud computing services. *Software Engineering Challenges of Cloud Computing, IEEE*, 2009:44–52.

Rozsnyai, S., Vecera, R., Schiefer, J., and Schatten, A. (2007). Event cloud - searching for correlated business events. *E-Commerce Technology, IEEE International Conference on, and Enterprise Computing, E-Commerce, and E-Services, IEEE International Conference on*, 0:409–420.

Santos, N., Gummadi, K. P., and Rodrigues, R. (2009). Towards trusted cloud computing. In *Proceedings of the 2009 conference on Hot topics in cloud computing*, HotCloud'09, pages 3–3, Berkeley, CA, USA. USENIX Association.

Schmidt, M. T., Hutchison, B., Lambros, P., and Phippen, R. (2005). The enterprise service bus: making service-oriented architecture real. *IBM Syst. J.*, 44(4):781–797.

Silveira, P., Rodriguez, C., Casati, F., Daniel, F., D'Andrea, V., Worledge, C., and Taheri, Z. (2010). On the design of compliance governance dashboards for effective compliance and audit management. In Dan, A., Gittler, F., and Toumani, F., editors, *Service-Oriented Computing. ICSOC/ServiceWave 2009 Workshops*, volume 6275 of *Lecture Notes in Computer Science*, pages 208–217. Springer Berlin / Heidelberg.

Skene, J. (2007). *Language support for service-level agreements for application-service provision*. PhD thesis, Department of Computer Science, UCL.

Sotto, L., Treacy, B., and McLellan, M. (2010). Privacy and data security risks in cloud. *Computing Electronic Commerce & Law Report*, 15:186.

Swanson, M., Bowen, P., Wohl Phillips, A., and Gallup, D., D. L. (2010). Contingency planning guide for federal information systems. *NIST Special Publication 800-34 Rev. 1*.

Weber, R. (1998). *Information Systems Control and Audit*. Pearson Education.

Wood, L. (2009). Cloud computing and compliance: Be careful up there. Computerworld.

Zhang, C., De Sterck, H., Aboulnaga, A., Djambazian, H., and Sladek, R. (2010). Case study of scientific data processing on a cloud using hadoop. In Mewhort, D., Cann, N., Slater, G., and Naughton, T., editors, *High Performance Computing Systems and Applications*, volume 5976 of *Lecture Notes in Computer Science*, pages 400–415. Springer Berlin / Heidelberg.