Jan Jürjens, Kurt Schneider (Hrsg.): Software Engineering 2017,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2017 115

Modellbasierte Datenschutzprüfung datenintensiver Cloud Dienste

Eric Schmieders[1], Andreas Metzger[1] und Klaus Pohl[1]

Abstract: Wir stellen einen Ansatz vor, der Cloud-Systeme hinsichtlich deren Einhaltung von Datenschutzrichtlinien zur Laufzeit überwacht.

Keywords: Cloud Computing, Datenschutz, Laufzeitmodelle.

1 Einleitung

Datenschutzrichtlinien, wie die Datenschutz-Grundverordnung der EU, geben strenge Richtlinien für die Speicherung und Verarbeitung personenbezogener Daten vor. So dürfen personenbezogene Daten prinzipiell nur innerhalb der EU-Grenzen gespeichert und verarbeitet werden. Eine Übertragung der Daten in Nicht-EU-Länder setzt voraus, dass diese Länder ausreichende Datenschutzmechanismen gewährleisten.

Werden Cloud-Systeme zur Datenverarbeitung eingesetzt, dann können aufgrund der Cloud-Elastizität Software-Komponenten dynamisch – auch zwischen Rechenzentren – verschoben werden (sog. „Replikation“ und „Migration“). Diese Verschiebungen sind zur Entwurfszeit nicht bekannt, da sie erst zur Laufzeit zur Optimierung der Performanz, der Verfügbarkeit und der Kosten ausgelöst werden. Durch die dynamische Verschiebung von Komponenten, wie z.B. von Datenbanken oder Analysemodulen, könnten Daten in unerlaubte Länder gelangen.

Unser Beitrag (siehe [SMP15]) ist der R-PRIS-Ansatz, der Cloud-Systeme zur Laufzeit bezüglich der Einhaltung von Datenschutzrichtlinien überwacht. Abbildung 1 zeigt die wesentlichen beiden Schritte des R-RPRIS Ansatzes:

1. Monitoring-Informationen über Verteilung und geographische Lage von Software-Komponenten werden gemeinsam mit Architekturwissen aus der Entwurfszeit zu einem Architektur-Laufzeitmodell verdichtet. Ein R-RPRIS Laufzeitmodell umfasst Informationen über Software-Komponenten, Aufrufbeziehungen zwischen diesen Komponenten und Geolokationen.
2. Eine Graph-Analyse auf dem jeweils aktuellen Laufzeitmodell für ein Cloud-System zeigt an, ob Daten in unerlaubte Länder gelangen können, und liefert somit Hinweise auf potentielle Datenschutzverletzungen.

[1] paluno (The Ruhr Institute for Software Technology), Universität Duisburg-Essen, 45127 Essen,
<vorname>.<nachname>@paluno.uni-due.de

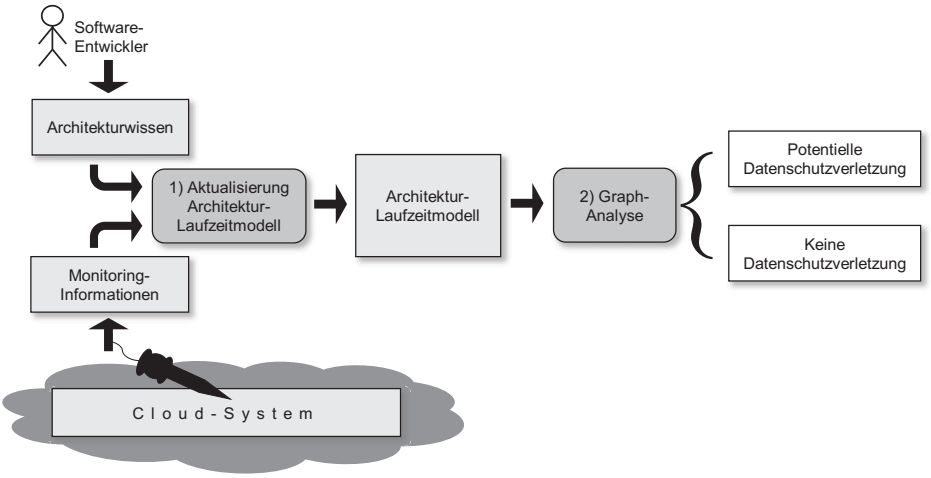

Abb. 1: Der R-PRIS-Ansatz zur Datenschutzüberwachung von Cloud-Systemen

Bisherige Lösungen zur Überwachung von Datenschutzrichtlinien adressieren nicht die besonderen Herausforderungen datenintensiver Cloud-Anwendungen:

- *Skalierbarkeit:* Moderne Datenanalyseanwendungen (z.B. hadoop) bestehen aus einer sehr hohen Zahl an Komponenten. R-PRIS ermöglicht die performante Überprüfung der Datenschutzrichtlinien auch bei großen Laufzeitmodellen.
- *Differenzierung:* Nicht alle Daten sind personenbezogen. So sind Kundendaten personenbezogen, eine anonymisierte Verkaufsstatistik nicht. R-RPRIS repräsentiert notwendige Informationen zur Differenzierung der Daten im R-RPRIS-Laufzeitmodell und berücksichtigt diese bei der Graph-Analyse.

Die R-RPRIS-Technik entstand im iObserve-Projekt des DFG-Schwerpunktprogramms SPP1593 (“Design For Future – Managed Software Evolution”).

Literaturverzeichnis

[SMP15] E. Schmieders, A. Metzger, K. Pohl, “Runtime model-based privacy checks of big data cloud services,” in 13th Int’l Conference on Service Oriented Computing (ICSOC2015), Ghoa, India, November 16-19, 2015, ser. Lecture Notes in Computer Science, A. Barros, D. Grigori, N. C. Narendra, H. K. Dam, Eds., vol. 9435. Springer, 2015