

Datenintegration mittels der regelbasierten Replikationsstrategie RegRess

Heiko Niemann

Lieken IT Service GmbH, heiko.niemann@lieken-it.de

Wilhelm Hasselbring

Universität zu Kiel, Software Engineering, wha@informatik.uni-kiel.de

Wenn Daten redundant in heterogenen Systemen gespeichert sind, dann wird für die Integration der Daten eine geeignete Replikationsstrategie benötigt. Die hier vorgestellte regelbasierte Replikationsstrategie RegRess erlaubt eine Konfiguration durch Replikationsregeln und eine Adaption zur Laufzeit, weil bei der Inferenz, durch die Schreib- und Lesezugriffe koordiniert werden, Systemzustände berücksichtigt werden. Dadurch kann für ein Anwendungsszenario hinsichtlich der konkurrierenden Ziele Konsistenz, Verfügbarkeit und Performance ein optimaler Kompromiss erreicht werden.

Die Integration von Informationssystemen in heterogenen Systemlandschaften bedeutet häufig ein Abgleich von Daten lokaler, autonomer Systeme. Als Beispiel dienen klinische Informationssysteme, die im Allgemeinen aus mehreren heterogenen Systemen bestehen, wobei jedes System Patientendaten speichert. Einerseits soll nun die Datenmanipulation auf den verschiedenen Systemen möglichst unter Wahrung der Konsistenz erfolgen, d.h. die Datenspeicherung muss entweder zentral erfolgen oder die Daten werden quasi zeitgleich abgeglichen. Andererseits sollen die beteiligten Systeme im Allgemeinen ihre Autonomie beibehalten, was eine lokale Speicherung erfordert und einen asynchronen Abgleich bedingt, um die Verfügbarkeit und Performance nicht bzw. nicht zu stark einzuschränken. Für derartige Anforderungen werden geeignete Replikationsstrategien benötigt, die einen optimalen Kompromiss hinsichtlich der Ziele der Replikation bieten, nämlich hinsichtlich der Konsistenz, der Verfügbarkeit und der Performance. Hierbei handelt es sich um konkurrierende Ziele, d.h. die Verbesserung eines dieser Ziele bedingt im Allgemeinen eine Verschlechterung der anderen Ziele. Beispielsweise ist eine Verbesserung der Konsistenz häufig nur zu Lasten der Verfügbarkeit und der Performance zu erreichen oder eine Erhöhung der Performance ist nur dann möglich, wenn Abstriche bei der Verfügbarkeit und der Konsistenz hingenommen werden.

In dieser Arbeit wird die regelbasierte Replikationsstrategie RegRess vorgestellt, die dazu beiträgt, für ein gegebenes Anwendungsszenario einen möglichst optimalen Trade-Off bezüglich der Replikationsziele zu erreichen. RegRess wird durch die Vorgabe einer Menge von anwendungsbezogenen Replikationsregeln konfiguriert. Vor jedem Schreib- oder Lesezugriff werden mittels Inferenz der Regeln die betroffenen Replikate bzw. die Art des Zugriffs bestimmt. Weil bei der Inferenz protokollierte Systemzustände berücksichtigt werden, passt sich RegRess zur Laufzeit dem Systemverhalten an.