Semantisch beschriebene Grid-Dienste in einem Semantik Webservice Framework

Dirk Mühlenberg & Sandro Leuchter

Abt. Interoperabilität und Assistenzsysteme
Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung IOSB
Fraunhoferstraße 1
76131 Karlsruhe
{vorname.nachname}@iosb.fraunhofer.de

Abstract: Grid-Middleware erlaubt den verteilten Zugriff auf Software- und Hardwareressourcen, indem sie prinzipiell solche Ressourcen beschreibt und einem Anwender über Protokoll und API auf seinem Arbeitsplatz zur Verfügung stellt. Ressourcen liegen auf entfernten Knoten, ihr Zustand ist jederzeit ermittelbar und Daten werden über definierte Protokolle ausgetauscht. Höherwertige Anwendungen werden durch Aggregation einzelner Ressourcen in Workflows komponiert, die durch entsprechende Werkzeuge vom Anwender modellierbar sind. Im Bereich „Kooperative Informationssysteme“ auf Basis einer SOA benötigt man aber eine Architektur, die es dem Anwender erlaubt, durch Formulierung eines Informationsinteresses einen dynamischen evtl. zyklischen Workflow zu starten, der zum Anfragezeitpunkt noch gar nicht bekannt ist, sondern vom System durch Vergleich der verfügbaren Ressourcen mit den angefragten Informationsobjekten erzeugt wird. Statische Workflowsysteme z.B. auf Basis von WS-BPEL (Web Services Business Process Execution Language) eignen sich nicht für diese Anforderung. Stattdessen verwenden wir einen Ansatz aus dem Bereich semantische Webservices, um zielgetriebene verteilte Aufgabenverarbeitung in einem Anwendungsbereich kooperativer Informationssysteme für die Luftbildauswertung zu realisieren. Jedoch berücksichtigen die Standardframeworks aus diesem Bereich keine Grid-Dienste und die dort eingeführten Formate, eine Integration ist aufgrund derselben technischen Schnittstelle (WSRF – Web Service Resource Framework) aber möglich. Die Hauptaufgabe dafür liegt in der Integration der Beschreibungselemente der Grid-Ressourcen in den Erzeugungsprozess des Workflows als nicht-funktionale Parameter. Die hier vorgestellten Arbeiten sind Bestandteil zu einem möglichen Semantik Grid Ansatz als Basis für eine service-orientierte Architektur.

1 Einleitung

Moderne Informationssysteme mit verteilten Ressourcen wie Datenbanken und Fachanwendungen erlauben es heutzutage dem Nutzer in einem Portal transparent auf Daten und Dienste zuzugreifen ohne genaues Wissen über den Ort der Daten oder der Anwendungen voraus zu setzen. Wird eine service-orientierte Architektur (SOA) basierend auf Webservices genutzt, so gibt es unterschiedliche Ansätze, um Ressourcen in einem Netz zu beschreiben. Die Beschreibung von IT-Ressourcen ist gerade im Konzept eines Software-Grids ein entscheidendes Basiselement, Grid-Ressourcen werden anhand wohl definierter Parameter beschrieben und zur Laufzeit durch entsprechende Dienste (sogenannte Information Provider) veröffentlicht und so dauerhaft Konsumenten zur Auswertung zur Verfügung gestellt. Diese Beschreibungen sind jedoch ausgelegt auf die Bedürfnisse von naturwissenschaftlichen wie simulationslastigen Anwendungen, d.h. Parameter zur Beschreibung von Hardware, Betriebssystem und Laufzeitgrößen spielen hier die wesentliche Rolle, um entsprechende Jobs an ausgesuchte Knoten im Grid zu delegieren. Die vorhandenen Beschreibungen können aber prinzipiell um anwendungsnahe Elemente erweitert werden. Die Nutzung von Ontologien ermöglicht die Ressourcen hierfür semantisch zu beschreiben.

2 Semantische Beschreibung von Grid-Ressourcen

Ein Framework, das sich zur Ontologie-basierten Dienstbeschreibung anbietet, ist das „Semantic Metadata Discovery and Monitoring System“ (S-MDS), das im Rahmen des AIST-GTRC Database Grids entwickelt wurde [MK06]. Diese Software erlaubt es, in einem Globus Toolkit (GT4) Grid Dienste mit Metadaten in Form von Ontologieinstanzen zu versehen. Die Informationen werden zur Laufzeit über GT4 Standardmechanismen (Monitoring und Discovery Service - MDS) gesammelt und auf diese Ontologie-Instanzen abgebildet. Die Beschreibungen und ihre Zuordnungen können mit Editoren erstellt und im GT4 standardkonform über Information Provider veröffentlicht werden. Eine Untermenge der Metadaten ist dann z.B. über eine SPARQL Anfrage von Konsumenten zu ermitteln. Dieses Konzept ist für eine Integration in eine semantische Grid-basierte SOA besonders geeignet, da es konform ist mit den GT Mechanismen und eine flexible und gut erweiterbare Basis für die Modellierung von Dienstbeschreibungen darstellt.

2.1 Dienstbeschreibung mittels Ontologien als nicht-funktionale Parameter

Die Anforderungen an einen Fachdienst zerfallen in funktionale und nicht-funktionale Anteile. Der funktionale Anteil beschreibt die benötigten und die zurückgelieferten Daten sowie Seiteneffekte und werden in einer technischen Schnittstelle realisiert. Der nicht-funktionale Anteil ist in seiner Beschreibung davon unabhängig und spiegelt die Eigenschaften eines Dienstes wider, die es z.B. erlauben, einen Dienst aufzufinden oder Kosten, Qualität und seinen Nutzen zu bewerten. Solche Beschreibungen können unabhängig von einer Ausführung des Dienstes sein (wie etwa die Beschreibung der Fähigkeiten eines Detektionsverfahrens) oder aber von der Ausführung abhängen, wie

etwa Payload-abhängige Tarife. Erst die nicht-funktionalen Parameter ermöglichen die korrekte und effiziente Nutzung von Diensten, etwa gesteuert durch Präferenzen in einem Nutzerprofil über eine aufgabenangepasste Kosten-Nutzen-Rechnung oder aber die Qualitätsbewertung eines Dienstes durch Einzel- oder Gemeinschaftsbewertung der Nutzer im Netz. Die Repräsentation der Parameter als Wissensbasis erlaubt zudem regelbasiert zu arbeiten und über die verschiedenen Informationen Schlussfolgerungen zu ziehen.

Als Basisontologie muss im Sinne der Kompatibilität im GT4 Umfeld ein Modell basierend auf GLUE-CE (Grid Labatory for a Uniform Environment) [An09] gewählt werden. „CE“ steht für Computing Element ein Basiskonzept für eine Entität, die eine Grid-Ressource managt. Diese Basis wurde im Rahmen eines Experimentalsystems für die Anwendung Luftbildauswertung erweitert um Anteile, die auf den Grundkonzepten „anwendungsrelevantes Objekt“ (ARO), „Information“, „Dienst“, „Experte“ und „Nutzereigenschaften“ basieren. Die Fähigkeiten eines Experten oder Dienstes wurden aus Netzen von AROs modelliert, die Güte oder Qualität aus dem Übereinstimmungsgrad der Fähigkeit mit der geforderten Fähigkeit aus der Nutzeranfrage ableiten. Der funktionale Anteil der Schnittstelle wird im Webservice/Grid Umfeld in einer WSDL (Web Service Description Language) Datei gemäß dem zugehörigen W3C Standard beschrieben, der nichtfunktionale Anteil wird im nächsten Abschnitt ausgeführt.

2.2 Umsetzung in S-MDS

S-MDS erlaubt es, Standard Grid Dienste mittels einer ontologiebasierten Beschreibung (Metadaten) zu hinterlegen. Die Ontologie ist in OWL (Web Ontology Language) zu formulieren. Eine Ausgangsform basierend auf GLUE-CE ist in OWL verfügbar, die mit einem Ontologieeditor wie Protégé zu erweitern ist. Die eigentlichen Laufzeitinformationen der nicht-funktionalen Eigenschaften werden durch einen Grid Resource Provider gestellt, der je nach Anwendung (also Satz von Dienstparametern) zu implementieren ist. Für die Standard Grid-Parameter gibt es bereits vorgefertigte Klassen. Hat man die anwendungsspezifische Ontologie erstellt, so lassen sich anhand eines festgelegten Ablaufs in einem Editor die für einen Dienst wichtigen Konzepte mit ihren Werten verbinden. Die derzeitige Implementierung von S-MDS lässt als Zuordnung von Parameter und Wert derzeit nur XPath (XML Path Language) Ausdrücke [Be07] zu. Die Werte müssen dabei einer XML Datei entstammen, die vom eigentlichen Wertelieferant (Resource Provider) zu befüllen ist. Hervorzuheben ist die Trennung von Metadatenlieferant und der eigentlichen Darstellung der Metadaten als Ontologieinstanzen innerhalb von S-MDS. Somit ist es prinzipiell dem Anbieter eines Dienstes im Grid überlassen, wie er seine Metadaten Grid-konform präsentiert, S-MDS sorgt für die Übersetzung in eine Semantik Web konforme Darstellung und die zugehörige Zugriffsmöglichkeit.

3 Zielgetriebene Problemlösung für Web Services

Als Middleware für eine zielgetriebene Problemlösung wurde das WSMO (Web Service Modelling Ontology) Framework [Ro06] gewählt, da es basierend auf Standards im Bereich Semantik Web sowie W3C Webservices entwickelt wurde und durch seinen flexiblen modularen Aufbau eine leichte Erweiterungsmöglichkeit für zusätzliche Dienstschnittstellen und -protokolle bietet. WSMO erlaubt es, Webservices mit OWL-konformen Ontologien zu beschreiben. Das Framework besteht aus der Laufzeitumgebung WSMX (Web Service Modelling eXecution environment), auf der Webservices, die mittels WSML dargestellt sind, ausgeführt werden. Es gibt zwei Werkzeuge zur Modellierung: „Studio“ und „Toolkit“. Mit beiden können Dienste, Ontologien und Workflows entwickelt, getestet und in Repositories veröffentlicht werden. Webservices in WSMO können anhand ihrer Beschreibungen gesucht und für eine bestimmte Aufgabe ausgewählt werden (Matchmaking). Die eigentliche Stärke von WSMO liegt aber in der Art und Weise, wie Workflows ausgeführt werden können. Zum einen gibt es die Möglichkeit, Workflows zu designen oder auch aus anderen Darstellungsformen (z.B. EPML – Event-driven Process Chain Model Markup Language in ARIS) zu importieren, zum anderen aber kann man ein Ziel angeben, zu dem es keinen Workflow gibt. Die Laufzeitumgebung versucht anhand der zur Verfügung stehenden Ressourcen (Dienste und Abbildungsvorschriften), dieses Ziel zu erreichen. Hierbei müssen evtl. komplexe Aufgaben erledigt werden. Beispiele sind das Mapping von unterschiedlichen Ontologien aufeinander, das Auswerten komplexer Choreografien oder das Mapping unterschiedlicher Protokolle (durch Prozessabbildung innerhalb der Choreografie) aufeinander. Ein Ziel (Goal) kann einfach die Erzeugung einer Ontologie-Instanz (im einfachsten Fall über den Aufruf eines Dienstes) in einer Wissensbasis oder die Auswertung eines komplizierten logischen Ausdrucks sein.

3.1 Integration der Grid Dienste in WSMO

Ein Webservice in WSMO ist eine Abstraktion von der eigentlichen darunterliegenden Implementierung, die z.B. als W3C Web Service vorliegt. Die Servicebeschreibung wird in WSML [Ro06] formuliert, zur Ausführung des Services muss der abstrakt beschriebene Dienst und seine Eingangsdaten einem Grounding unterzogen werden, d.h. die funktionalen und nicht-funktionalen Eigenschaften, die als Ontologieinstanzen in WSMO vorliegen, werden durch ein *Lowering*[1] in XML Instanzen gemäß der Spezifikation unterzogen. Resultate werden durch ein entsprechendes *Lifting* in Ontologieinstanzen gewandelt. Dieser Prozess ist abgetrennt in einem eigenen Modul verfügbar und für jeden Service zu implementieren. Das *Lowering* muss ausprogrammiert werden, das *Lifting* lässt sich durch ein XSL Transformation darstellen. Ein Grid-Dienst ist ebenfalls ein Webservice, das WSRF basierte Interface wird jedoch über eine eigene Grid-konforme Schnittstelle innerhalb der *Lowering-Lifting*-Komponente (LILO) angesprochen, da z.B. im Grid die Sicherheitsanforderungen voraussetzen, dass der Anfragende authentifiziert ist und eine Authorisierung für die angefragte Ressource besitzt.

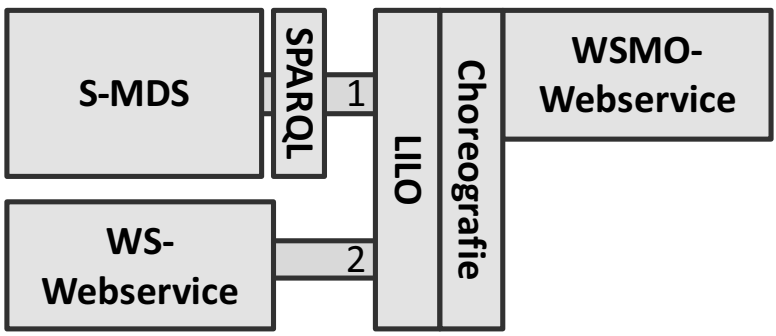

Abbildung 1: Integration von S-MDS in WSMO über Choreografie

[1] Lowering ist der Übergang (*Grounding*) vom Domänenmodell (in WSML) nach SOAP (in XML), Lifting der umgekehrte Vorgang.

3.2 Nutzung der Dienstparameter in WSMO

Die Dienstparameter (Metadaten) können auf der Gridseite über einen eigenen Metadatendienst angefragt werden, der sich auf den S-MDS Services abstützt. Der Konsument muss den interessierenden Dienst benennen und erhält die gewünschten Metadaten zurück. Dies ist auf der WSMO Seite über einen eigenen Webservice realisiert, der schließlich den Metadaten Grid Dienst aufruft. Somit lässt sich in der Choreografie von WSMO Services ein Ziel angeben, das die Metadaten zu dem eigentlichen Grid Dienst bereitstellt (Schritt 1 in Abbildung 1**Fehler! Verweisquelle konnte nicht gefunden werden.**) und mittels Schlussfolgerung oder anderer Mechanismen das Auffinden, die Choreografie oder Orchestrierung von Diensten innerhalb von WSMO beeinflusst (Schritt 2 in Abbildung 1). Insbesondere bei Trennung von in funktionaler Hinsicht gleichartiger Dienste durch nicht-funktionale Parameter wie Fähigkeitsbeschreibungen (z.B. bei Verfahren zur Objektidentifikation) liegt der Wert der Repräsentation eines solchen Parameters durch Ontologien in der Mächtigkeit der Beschreibung und Möglichkeit zur logischen Verarbeitung der Wissensbausteine bzgl. der Anforderungen an diesen Parameter (wie Detektion von Fahrzeugen im Gegensatz zu Detektion von Personen) durch die Nutzeranfrage.

4 Verwandte Arbeiten

In [Pa06] ist eine Übersicht über Semantik Grid Ökosysteme. Die Autoren bemühen ein Softwaremuster der ‚Cooperative Semantic Aware and Semantic Provisioning Services' sowie den Methodenansatz der evolutionären bzw. der kompletten Neuentwicklung. In dieser Arbeit werden die gewählten Ansätze einer Analyse bzgl. eines umfangreichen Kriterienkatalogs unterzogen, sowie in einem weiteren Kapitel näher beschrieben.

Ein Ansatz, der von den gewählten Frameworks (WSMO + Grid) dem hier vorgestellten am nächsten kommt, ist in [Sh09] zu finden. Das Konzept und der Anspruch sind jedoch verschieden, der Autor erweitert WSMO (und somit WSML) um Konzepte für Grid-Services (sind WSMO Web-Services) und Jobs (sind WSMO Goals) und erläutert anhand eines Testfalls die Interaktion von WSMO mit dem darunterliegenden Grid-Framework. Der Schwerpunkt in [Sh09] liegt in der Spracherweiterung und den dafür notwendigen Erweiterungen der Laufzeitumgebung WSMX und nicht in der semantischen Annotation der Ressourcen und der Nutzung dieser Metadaten in WSMO. Außerdem war es wichtig, die verwendeten Frameworks unverändert zu lassen, um neue Versionen leichter nutzen zu können.

5 Zusammenfassung und Ausblick

Im Rahmen des hier beschriebenen experimentellen kooperativen Informationssystems zur Unterstützung der Luftbildauswertung wurden Beiträgen zu einem Semantik Grid Ansatz gezeigt, mit denen anwendungsnahe Metadaten von Grid Diensten bei zielgetriebenen Problemlösungen einbezogen werden können. Ein Schwerpunkt der

Arbeiten lag in der Wiederverwendung bestehender Konzepte zu kooperativen Informationssystemen in einer neuen Architektur, die auf Standards in den Bereichen Semantik Web und Software-Grid beruht. Ebenso sollten die verwendeten Frameworks keinerlei Veränderungen und Einschränkungen unterzogen werden, um Migration und vollen Funktionsumfang zu garantieren. Eine Besonderheit der hier beschriebenen Integration ist, dass Grid Dienste und Semantic Web Services gleichermaßen genutzt werden können.

Zukünftig steht die Integration dieser Anteile in den Semantic Service Bus [Ka07] an, wodurch sogar die Einbindung weiterer Protokolle und Programmierschnittstellen ermöglicht wird.

6 Literaturverzeichnis

[An09] Sergio Andreozzi et.al: Glue specification v. 2.0. Technical report, Open Grid Forum, March 2009.

[Be07] Anders Berglund et.al: XML Path Language (XPath) Version 2.0, W3C Recommendation, January 2007.

[Ka07] Dimka Karastoyanova et.al: Semantic Service Bus: Architecture and Implementation of a Next Generation Middleware. ICDE Workshops 2007: 347-354.

[MK06] S.Mirza and I.Kojima, "S-MDS: A Semantic Information Service for Advanced Resource Discovery and Monitoring in WS-Resource Framework", GGF16 3rd Semantic Grid Workshop, Feb., 2006.

[Pa06] Pinar Alper et al: Semantic Aware Grid Services, Deliverable D1.4, OntoGrid project, August 2006.

[Ro06] D. Roman et.al: WWW: WSMO, WSML, and WSMX in a nutshell. In Proceedings of the first Asian Semantic Web Conference (ASWC 2006), Beijing, China, September 3-7, 2006.

[Sh09] M. Omair Shafiq: Semantic Grid: Extending Semantic Web, Master Thesis in Computer Science, March 2009, University of Innsbruck.