

» SAP INFO deutsch: Home » Technologie

Artikel Autoren

August 09, 2006 || Fraunhofer FIRST verbindet SOAs und Grid-Computing:

Zwei Konzepte kommen sich näher (Teil 2)

In der Verbindung von serviceorientierten Architekturen (SOAs) und Grid-Computing besteht beträchtliches Potenzial für künftige IT-Lösungen. Während Grid-Technologien einen organisationsübergreifend einheitlichen Zugriff auf verteilte IT-Ressourcen ermöglichen, liefern SOAs den technischen Rahmen und den methodischen Überbau dafür. In Teil 1 unseres Artikels wurden die Grundlagen beider Konzepte erläutert. Der zweite Teil verdeutlicht dies nun anhand ausgewählter Projekte.

Teil 1

Das Fraunhofer Institut für Rechnerarchitektur und Softwaretechnik (Fraunhofer FIRST) ist an einer Vielzahl von Forschungs- und Industrie-Projekten zur Verbindung von SOA und Grid-Computing beteiligt. Als Mitbegründer der Fraunhofer Grid Alliance (FhGA) wirkt Fraunhofer FIRST maßgeblich am Aufbau des Fraunhofer Resource Grids (FhRG) mit. Diese institutsübergreifende Grid-Plattform der Fraunhofer Gesellschaft stellt umfangreiche IT-Ressourcen – Hardware, Daten, Software und Dienste – in einer SOA zur Verfügung.

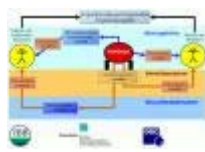
Die Projekte zur Verbindung von SOA mit Grid-Computing beschäftigten sich unter anderem mit der Fragestellung, wie sich geografisch verteilte Ressourcen effektiv in einem Arbeitsablauf kombinieren lassen. Hier hat Fraunhofer FIRST unter anderem den auf Petrinetzen basierenden Workflow-Dienst "Grid Workflow Execution Service (GWES)" entwickelt, mit dem sich vorhandene Web-Services, Grid-Services und gewöhnliche Programmaufrufe in einem automatisierten Arbeitsablauf auf einfache Art und Weise integrieren und ausführen lassen <http://www.gridworkflow.org/gwes>. Eine weitere wesentliche Aufgabe ist es, IT-Ressourcen anwenderfreundlich in einem serviceorientierten Grid bereitzustellen. Hier verfolgt Fraunhofer FIRST verschiedene Konzepte. Mit "Instant-Grid" hat das Institut eine Knoppix-basierte Live-CD realisiert, die aus einer vorhandenen Rechner-Infrastruktur innerhalb weniger Minuten eine funktionsfähige Grid-Umgebung macht <http://instant-grid.de/>. An diesem Projekt ist unter anderem die Gesellschaft für wissenschaftliche Datenverarbeitung mbH in Göttingen und die ed-media GmbH in Kaiserslautern beteiligt. Gemeinsam mit Anwendern und Entwicklern am Deutschen Zentrum für Luft- und Raumfahrt wurde zudem eine "VirtualLab-Plattform" entwickelt, die mit Hilfe von Spezialsoftware web- und gridbasiertes Arbeiten für die ganze Bandbreite des wissenschaftlichen und ingenieurtechnischen Rechnens möglich macht <http://www.first.fraunhofer.de/vlab>. Da für die Verbindung von SOA und Grid-Computing Sicherheitskonzepte notwendig sind, die eine zuverlässige Zugangskontrolle für Web-Services bieten und zugleich die Vertraulichkeit von Nutzerdaten gewährleisten, arbeitet Fraunhofer FIRST mit der Universität Potsdam auch am Prototyp einer feingranularen Grid- und Web-Service-Sicherheitslösung, die eine zertifikatsbasierte Zugriffskontrolle und die eingeschränkte Delegation von Rechten für Web-Services ermöglicht.

ERAMAS: Intelligentes Katastrophenmanagement

Die Vorteile des Zusammenspiels von Grid-Computing und serviceorientierten Architekturen lassen sich anhand von drei Praxisbeispielen erläutern.

Das Analyse- und Managementsystem ERAMAS (Environmental Risk Analysis and Management System) ermöglicht es, die Ausbreitung von Krebs erregenden und chemisch-toxischen Schadstoffen in der Luft, im Boden und im Grundwasser sowie den Gefährdungsgrad für den Menschen mit Hilfe hochwertiger Simulationsprogramme zu prognostizieren und zu bewerten. Einsatzmöglichkeiten sind Gefahrguttransporte, Störfälle in industriellen Anlagen oder terroristische Anschläge. Die Analyse erfolgt unter Echtzeitbedingungen, das heißt, die Rechenergebnisse müssen so schnell vorliegen, dass die Einsatzkräfte beispielsweise Evakuierungsentscheidungen treffen und umsetzen können, bevor die Schadstoffe sich ausbreiten können.

ERAMAS, das von Fraunhofer FIRST in Zusammenarbeit mit der Dresdner Grundwasser Consulting GmbH und dem Ingenieurbüro Beger für Umweltanalyse und Forschung entwickelt wurde, kann außerdem als kommerziell betriebenes Dienstleistungssystem genutzt werden – zur Vorfeldanalyse für Genehmigungsverfahren, für das Erstellen von Maßnahmeplänen oder die Schulung von Einsatzkräften. Der Prototyp des Systems lässt sich rasch an die Anforderungen der Kunden – etwa Chemieunternehmen oder Feuerwehren – anpassen oder erweitern.



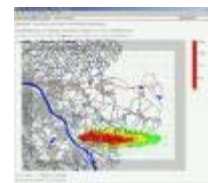
Modellkopplung in einer SOA

In ERAMAS kommt die Technologie des Fraunhofer Resource Grids (FhRG) zum Einsatz, das auf dem Globus-Toolkit basiert. Mit Hilfe von verschiedenen Diensten ermöglicht dieses Softwarepaket den einheitlichen Zugriff auf verteilte Ressourcen. Dank der FhRG-Technologie lassen sich die einzelnen bei den Projektpartnern lokalisierten Soft- und Hardwarekomponenten von ERAMAS – etwa die Simulationsmodelle für die Luft- und Bodenausbreitung von Schadstoffen oder die organisationsübergreifend verteilten Rechenkapazitäten – über das Internet in einer SOA koppeln. Dazu werden sie gemäß eines Komponentenmodells gekapselt. Spezielle Grid-Software überwacht die Hardwareressourcen, optimiert ihre Auslastung und stellt weitere Grid-Funktionalitäten als Web-Services zur Verfügung. Mögliche Arbeitsabläufe sind in

Form von Workflow-Beschreibungen verfügbar und lassen sich bei Bedarf mit Hilfe eines Workflow-Dienstes auf dem Grid zuverlässig ausführen.

Komplexe Simulationen in Minutenschnelle

Die Stärke des Systems liegt in seiner Flexibilität. Je nach Problemstellung wählt es die geeignetsten Modelle und Systeminformationen aus. So berücksichtigt es die Rahmenbedingungen eines bestimmten Umweltproblems zum Beispiel in Bezug auf das betroffene Gebiet oder die Art der ausgetretenen Substanz. Verdunsten bei einem Unfall große Mengen giftiger Stoffe, bestimmt ERAMAS zunächst die Wetterdaten für den Unfallort und installiert ein zweckmäßiges Verdunstungsmodell auf den verfügbaren Hochleistungsrechnern. Dieses ermittelt die wahrscheinlichste und die maximal zu erwartende Verdunstungsrate in Abhängigkeit der Stoffparameter, der Bewölkung, der Windgeschwindigkeit und der Temperatur. Gleichzeitig berechnet das System aus den Daten einer Wetterstation und den Höhenstrukturen (Orographie) des Gebietes ein dreidimensionales Windfeld. Mit der Verdunstungsrate und dem Windfeld wird ein Ausbreitungsmodell erstellt, das die Schadstoffkonzentrationen in jeder Phase und für alle Punkte des betroffenen Gebietes bestimmt. Diese Ergebnisse wiederum werden mit Modellen gekoppelt, die die Versickerung der Schadstoffe und ihren Transport im Grundwasser prognostizieren und die Auswirkungen auf die Umwelt und die Gesundheit der Menschen abschätzen.



Simulationsergebnisse

Dank der Grid-Technologie kann der gesamte Arbeitsablauf – von der Erfassung der Unfalldaten bis hin zur Bewertung der kurz- und langfristigen Gefährdung für Mensch und Umwelt – auf wenige Minuten verkürzt werden. Da ERAMAS als verteiltes System realisiert ist, dessen Modelle von den Entwicklern betreut werden, lassen sich wissenschaftliche oder technische Neuerungen, beispielsweise neue Simulationsgebiete und weitere Transportpfade, jederzeit leicht einbinden.

K-Wf Grid: Durchblick im Dschungel der Dienste

Das Grid der Zukunft wird eine große Vielzahl von Diensten umfassen, die teilweise kooperativ aufeinander aufbauen (orthogonale, einander ergänzende Dienste), aber teilweise auch gegeneinander konkurrieren (Dienste gleicher oder ähnlicher Funktionalität von verschiedenen Anbietern). Um eine derart komplexe Umgebung sinnvoll zu nutzen, bedarf es einer unterstützenden Infrastruktur. Hier setzt das Projekt K-Wf Grid (Knowledge-based Workflow System for Grid Applications) an, in dem Forschungseinrichtungen und Firmen aus verschiedenen Ländern der Europäischen Union unter der Leitung von Fraunhofer FIRST zusammenarbeiten. Entwicklungsziel ist ein Expertensystem, das Wissen über das Grid sammelt und bereitstellt und beispielsweise die Frage beantwortet, welcher Workflow am häufigsten erfolgreich zur Lösung eines Problems verwandt wurde.



Architektur des K-Wf Grid

K-Wf Grid unterstützt die Anwender bei der Auswahl von Diensten und beim Erstellen komplexer Workflows für Grid-Anwendungen in einer serviceorientierten Architektur. Softwarekomponenten, deren Funktionalitäten als Dienste im Grid zur Verfügung stehen, können in unterschiedlichen Zusammenhängen zu größeren Anwendungen kombiniert und effektiv ausgeführt werden. Dazu überwacht das System sämtliche Interaktionen, etwa die Ausführung von Anwendungen auf dem Grid, und wertet sie aus. So steht eine breite Datenbasis zur Verfügung, die sich bei der Entwicklung von Workflows nutzen lässt und die Qualität der Services gewährleistet. Die Ausführungsdauer, die Zuverlässigkeit und die Qualität der Ergebnisse einer gekoppelten Anwendung lassen sich somit schon im Vorfeld abschätzen und optimieren.

Das System wird derzeit um spezifisches Anwendungswissen erweitert. Dazu zählen die Bedingungen, unter denen sich bestimmte Programme einsetzen und koppeln lassen oder die Bedeutung und der Gültigkeitsbereich von Eingabedaten. Auf Basis dieses Wissens ist das System dann in der Lage, den Anwendern Vorschläge für bewährte Arbeitsabläufe zu machen.



Workflows

Dass das K-Wf Grid funktioniert, beweisen drei verschiedene Anwendungen: Zusammen mit dem "Slovak Water Management Enterprise" werden Simulationen zur Flutvorhersage realisiert, bei denen – ähnlich wie bei ERAMAS – verschiedene Umweltsimulationsprogramme als Dienste zur Verfügung gestellt und zu einem Workflow kombiniert werden. Zweitens wurde unter Verwendung von K-Wf Grid für die italienische Stadt Genua ein Verkehrsmanagementsystem entwickelt, das die Auswirkungen von Sperrungen, Baustellen und Staus auf den Verkehrsfluss und die Umweltbelastung der Stadt prognostiziert. Drittens portierte die Firma LogicDIS S.A. ihr Client-Server-basiertes ERP-System auf das serviceorientierte K-Wf Grid, um die Ausführung von Prozessen in einem ERP-System zu optimieren.

MediGRID: Medizinische Grundlagenforschung

Das Verbundvorhaben MediGRID weist am Beispiel biomedizinischer Forschung nach, wie sich Grid-Dienste in der Medizin und in den Lebenswissenschaften anwenden lassen. Das Projektkonsortium unter der Leitung der Medizintechnik-Dachorganisation TMF e.V. besteht einerseits aus Vertretern der biomedizinischen Forschung wie den Universitäten und Universitätskliniken von Göttingen, Marburg, Konstanz, Dresden sowie der Berliner Charité. Andererseits bringen die Fraunhofer-Institute für Arbeitswirtschaft und Organisation (IAO) und FIRST sowie das Berliner Zuse-Zentrum ausgewiesene Kompetenz auf den Gebieten Grid-Computing und verteiltes Datenmanagement ein.

Die MediGRID-Pilotanwendung "RNAi Pipeline" soll funktionelle Eigenschaften von Proteinen mit unbekannter 3D-Struktur vorhersagen. Ausgehend von Proteinsequenzen werden anhand eines Vergleichs mit bekannten, klassifizierten Strukturen Proteinstrukturmodelle vorgeschlagen (fold recognition). Mit dieser Information analysiert die Anwendung in einem nächsten Schritt die Interaktion mit anderen Biomolekülen, so genannte "pathways". Da eine Vielzahl derartiger Sequenzen unabhängig zu untersuchen ist, kann die gesamte, aus drei Hauptkomponenten bestehende Verarbeitungs-Pipeline parallel über die Sequenzen durchlaufen werden.

Die bisherige Implementierung der Anwendung basierte bereits auf einer SOA – die Hauptkomponenten sind als statische

Web-Services realisiert und kommunizieren über SOAP. Damit ließ sich das oben dargestellte Potenzial an Parallelität jedoch noch nicht angemessen nutzen. Auch eine bloße Replikation der Dienste wird der Dynamik des Job-Aufkommens nicht gerecht und erfordert einen erheblichen Mehraufwand an Organisation und Administration. Im MediGRID-Projekt wird die Anwendung daher in ein auf Globus-Toolkit-4-basiertes Community-Grid für die biomedizinische Forschung in Deutschland integriert, um zusätzliche Rechenressourcen bedarfsgerecht einbinden zu können.

Dadurch dass die Komponenten der RNAi-Pipeline in Grid-Services migriert werden, lassen sich die benötigten Dienste vollkommen ressourcenschonend zuordnen. Die im Fraunhofer Resource Grid entwickelte Middleware und Ressourcenbeschreibung stellt sicher, dass aus der Masse der verfügbaren Rechenressourcen stets diejenigen ausgewählt werden, die für die Ausführung des Dienstes optimal geeignet und derzeit nicht voll ausgelastet sind.

Die Vorteile in Bezug auf Interoperabilität und Sicherheit, die dank der serviceorientierten Basisarchitektur bereits gegeben sind, werden durch die Grid-Integration um eine weitgehende Virtualisierung der Ressourcen ergänzt. Damit ist es möglich, die dem Anwendungsproblem inhärente Parallelität auf den verfügbaren Ressourcen optimal auszunutzen, ohne dass zusätzlicher Administrationsaufwand entsteht. In Entwicklung befindet sich derzeit ein vereinfachter webbasierter Zugang zu der Anwendung über das gemeinsame Grid-Portal des MediGRID-Projekts.

Artikel	Autoren
---------	---------



Michael John



Thilo Ernst



Andreas Hoheisel