

# Erste Schritte auf dem Weg zur Verarbeitung digitaler Geländemodelle mittels Grid Computing

Sandra Lanig, Arne Schilling, Beate Stollberg, Alexander Zipf

## Zusammenfassung

Digitale Geländemodelle (DGM) sowie 3-dimensionale Geodaten spielen eine wichtige Rolle in der Geographie. Zahlreiche Simulationen, wie z. B. Überflutungsmodellierungen und räumliche Analysen, benötigen sehr genaue Geländedaten. Während der Akquirierung dieser sehr hochauflösenden Daten, beispielsweise mittels Laserscanning-Verfahren, fallen aufgrund der hohen Punktdichte sehr große Datenmengen an. Klassische GIS können den Aufwand für die Verarbeitung und Analyse dieser massiven Geländerohdaten nicht befriedigend bewältigen. Um eine hohe Performanz sowie Speicherkapazität bei der Datenverarbeitung zu erreichen, erscheint Grid Computing als eine geeignete Alternative. Daher erfolgt die Entwicklung einer Reihe von Web Processing Services (WPS) zur Verarbeitung massiver Geodaten (insb. DGMs), welche als Grid Dienste verfügbar gemacht werden.

## 1 Einführung

Für die Verarbeitung großer Digitaler Geländemodelle (DGM) sind sehr rechenintensive geometrische Processing-Operationen notwendig. Mit dem Aufkommen des Grid Computing, welches die Rechenleistung ganzer Rechner-Cluster zur Verfügung stellen kann, bietet sich nun ein neuer Ansatz, um derartige Berechnungen in akzeptabler Zeit durchführen zu können. Hierbei darf der Begriff GRID nicht mit dem im GIS-Umfeld bekannten Rasterdatenformat verwechselt werden, sondern meint auf bestimmte Weise verbundene Rechner-Cluster. Noch sind viele Fragen der technischen Umsetzung offen, aber die Nutzung von Geoprocessing-Funktionen mittels einer Grid Infrastruktur stellt eine verlockende Perspektive für die Realisierung von sehr rechenintensiven Aufgaben in der Geodatenverarbeitung dar. Moderne Geoinformationsanwendungen werden zunehmend als Dienste-basierte Architekturen zur Erhöhung der Interoperabilität mittels standardisierter Schnittstellen realisiert. Derartige Geodateninfrastrukturen basieren auf den bekannten OGC Web Service (OWS) zur Datenhaltung, Suche und Visualisierung von Geodaten. Seit kurzem steht mit dem Web Processing Service (WPS) eine OGC Spezifikation zur Verfügung, die zudem das Verarbeiten der Geodaten ermöglicht (SCHUT 2007). In Vorarbeiten wurden schon mehrere dieser Processing-Funktionen als WPS realisiert (z.B. STOLLBERG & ZIPF 2007, 2008). In anderen Vorarbeiten wurden Java-Bibliotheken zur Verarbeitung von DGM-Daten realisiert (SCHILLING ET AL. 2007). Ziel des Projektes GDI-

Grid ([www.gdi-grid.de](http://www.gdi-grid.de)) ist es nun die beiden Welten Grid Computing und Geodateninfrastrukturen (GDI) zusammen zu bringen. Es geht dabei um die effektive und effiziente Erschließung sowie Prozessierung von Geodaten für mehrere Anwendungsszenarien (‘Geosimulation von Lärmausbreitung’, ‘Katastrophensimulation Hochwasser’ und ‘Routenoptimierung’). Hierbei werden auch Geländedaten in unterschiedlicher Auflösung und Skalierung benötigt. In diesem Beitrag wollen wir einen Teilbereich aus unseren Arbeitspaketen vorstellen. Es geht darum, die erwähnten DGM-Bibliotheken erstens funktional zu erweitern, sie zweitens als WPS-Prozesse zur Verfügung zu stellen und drittens diese WPS-Prozesse in die Grid Infrastruktur zu integrieren. Viertens sollen diese WPS DGM-Prozesse an ausgewählten Szenarien in eine Anwendung eingebunden werden, indem mehrere Dienste in einer Dienstekette einen Workflow abbilden. Zu den Möglichkeiten und aktuellen Problemen der Orchestrierung von OWS Diensteketten vgl. z.B. WEISER & ZIPF 2007, STOLLBERG & ZIPF 2007, KIEHLE ET AL. 2007.

## 1.1 Grid Computing

Der Begriff Grid Computing umfasst nach Forster (2002) drei Aspekte an technischen Fähigkeiten, welche ein Grid unterstützen sollte. Hierzu zählen:

- Koordinierung von Ressourcen, die nicht zentral vorgehalten werden,
- Nutzen standardisierten, offenen, universal Protokollen und Schnittstellen,
- Fördern nichttrivialer Dienstgüte (Quality of Service, QoS).

Oftmals wird der Begriff Grid mit einer Grid Middleware gleichgesetzt, welche jedoch lediglich ein Teilaspekt einer Grid Infrastruktur ist. Eine Grid Middleware bietet standardisierte, offene, universale Protokolle und Schnittstellen an, um Zugang zu räumlich verteilten Rechner- und Datenressourcen in einer übergangslosen und sicheren Art und Weise zu ermöglichen. Für die Standardisierung von Grid Middleware entwickelte das Global Grid Forum (GGF) (CZAJKOWSKI, K. ET AL. 2004) die Open Grid Service Architecture (OGSA), welches ein Konzept für die Entwicklung von offenen Grid Infrastrukturen basierend auf den Prinzipien der Service-orientierten Architektur (SOA) ist. Eine Referenzimplementierung dieses Konzeptes ist die Middleware Globus Toolkit 4 (GT4). Um Grid Anwendungen zu entwickeln, bietet GT4 eine Reihe von high-level Grid Infrastruktur Diensten, die Schnittstellen für das Verwalten von Rechenleistung, Speicher und andere Ressourcen beinhaltet. Hierdurch wird dem Nutzer ermöglicht, Grid-basierte Anwendungen zu entwickeln, ohne detailliert über das genutzte System bei der Verarbeitung eines Jobs im Grid oder den verteilten Speicherzugriffen bescheid zu wissen.

## 1.2 Prozessierung mittels Web Processing Service

Geodateninfrastrukturen (GDI) bieten Zugang zu global verteilten räumlichen Daten über standardisierte Web Services. Basierend auf interoperablen Schnittstellen des Open Geospatial Consortium (OGC) ermöglichen Web services das Verwalten, das Auffinden, das Verarbeiten und das Visualisieren von Geodaten. Im Dezember 2007 verabschiedete das OGC einen Dienst für die Verarbeitung bzw. Prozessierung von Daten mittels OGC basierten

Web service, den so genannten Web Processing Service (WPS). Dieser wurde entwickelt, um jegliche Art von GIS-spezifischen Verarbeitungsfunktionalitäten mittels Web Service bereitzustellen.

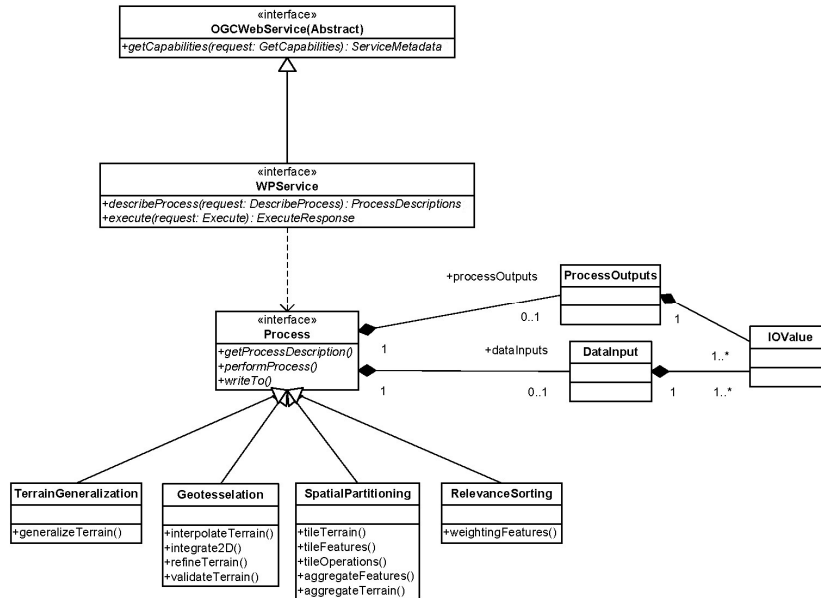


Abb. 1: WPS Schnittstelle - UML Diagramm mit implementierten Prozessen

Um verschiedene Geländeverarbeitungsfunktionalitäten basierend auf OGC Web Services anzubieten, erfolgt die Implementierung verschiedener WPS Geländeprozesse. Konkret handelt es sich bei den Diensten (vgl. Abb. 1) um Geländeverarbeitungsprozesse für die

- *Geotessellation* (WPS Generic Geotessellation Service Framework),
- *Parzellierung* (WPS Spatial Partitioning Service),
- *Generalisierung* (WPS Terrain Generalization Service),
- *Sortierung* (WPS Relevance Sorting Service).

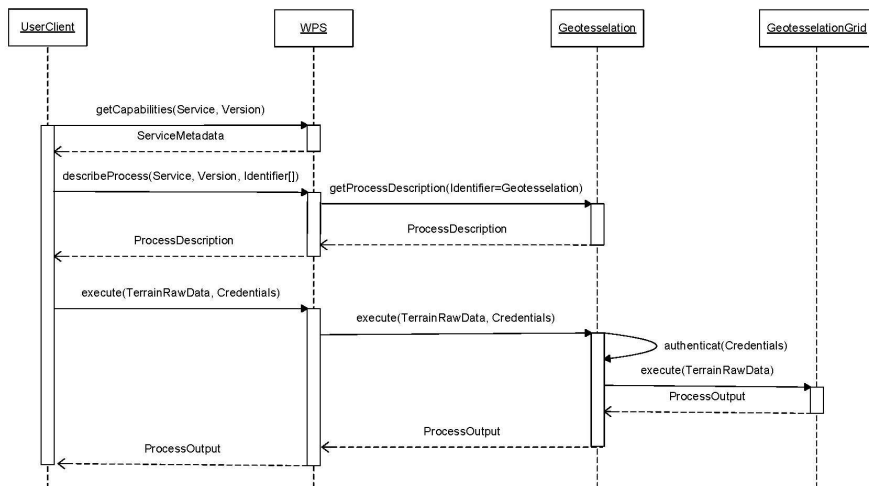
Generell werden im Grid Computing und bei OWS unterschiedliche Schnittstellen genutzt. Diese Inkompatibilitätsprobleme gilt es zu überbrücken. Normalerweise erfolgt die Ausführung des OGC WPS auf einem Application Server. Um die Prozessierungsdienste in das Grid zu transportieren, muss der WPS Grid-fähig gemacht werden. Dies erfolgt in einem eigenen Arbeitspaket im Rahmen des GDI-Grid Projektes. Die von uns entwickelten Processing-Algorithmen werden in dieses Framework eingebunden und gridifiziert.

## 2 Gelande-Prozessierung mittels Grid Computing

Im nachfolgenden Abschnitt erfolgt die Darstellung der einzelnen Gelandeverarbeitungs-Services. Zunachst erweitern wir die Funktionalitaten unserer bestehenden Gelandebibliothek. Danach werden die entwickelten Dienste als WPS verfugbar gemacht und die WPS Prozesse in eine Grid Infrastruktur mittels GT4 integriert. Zukunftig erfolgt die Integration der entwickelten WPS Gelandeoperationen innerhalb ausgewahlter Anwendungsszenarien mit weiteren Services als Workflow in einer Dienstekette.

### WPS Generic Geotessellation Service Framework

Der entwickelte Geotessellation Dienst generiert DGMs aus massiven Gelanderohdaten, z.B. Laserscanner-Daten. Die resultierende Gelandegeometrie wird als Triangulated Irregular Network (TIN) gespeichert. Die standardmaig implementierte Tesselerung basiert auf dem Delaunay Algorithmus. Diese Art der Tesselerung teilt den 2D-Raum in Dreiecke mit speziellen Bedingungen, so dass das resultierende TIN so gleichmaig als moglich ist.



**Abb. 2** Grid-fahiger WPS Geotessellation Service - UML Sequenzdiagramm

In Abb. 2 ist die interne Prozesskette der Tesselationsfunktionalitat in einem UML Sequenzdiagramm dargestellt. Ein Service Client nutzt das konventionelle OGC WPS Interface. Bei einer gewohnlichen GetCapabilities()- oder DescribeProcess()-Anfrage antwortet der WPS direkt mit einem Service-Metadatenokument bzw. mit der spezifischeren Prozessbeschreibung. Erfolgt eine Anfrage bzgl. Ausfuhrung des Prozesses mittels Execute()-Operation, so werden mit der Anfrage das Zertifikat (Credentials) fur die Identifikation und Authentifizierung des Nutzers sowie die Gelandedaten ubergeben. Falls die Authentifizierung erfolgreich ist, delegiert der OGC konforme Geotessellation WPS mittels Execute()-

Operation zu einem entsprechenden grid-fähigen WPS Geotessellation. Dieser zerlegt den Prozess in viele kleine Teilprozesse und sendet diese an den Job Scheduler Globus WS GRAM (Globus Web Service Grid Resource Allocator). Dieser übernimmt die Verteilung der einzelnen Jobs auf die jeweiligen Arbeitsknoten. Dort erfolgt die Geotessellierung.

### **WPS Spatial Partitioning Service**

Um ein verteiltes Verarbeiten der enormen Datenmengen in einer Grid Umgebung zu ermöglichen, müssen sowohl die Prozesse als auch die Daten parallelisiert werden. Die Parallelisierung der einzelnen Prozesse erfolgt in den einzelnen spezialisierten WPSen. Die räumliche Kachelung der Ausgangsdaten in kleinere Teilgebiete erfolgt in einem Vorverarbeitungsschritt mittels des Partitionierungsdienst. Abhängig von der Funktionalität bzw. der Datenstruktur sind verschiedene Partitionierungsartypen möglich, z.B. mittels einfacher Kachelung (Tiling) oder anhand von Geoobjekten. Insbesondere für DGMs eignet sich eine rechteckige Kachelung. Oft sind lediglich kleinere Teilgebiete aus einem Gesamtdatensatz relevant, welche z.B. durch administrative Grenzen gerastert werden können.

### **WPS Relevance Sorting Service**

In Abhängigkeit von anwendungsbezogenen Randbedingungen, wie z.B. die Bandbreitenbeschränkung beim Streaming oder die Leistungsbeschränkung aus der Sicht des Clients, machen es erforderlich, dass gewisse Geodatenobjekte oder Gebiete mit höherer Priorität verarbeitet bzw. an den Client bzw. nachfolgenden Prozessierungsschritt geliefert werden. Daher ist in einem Vorverarbeitungsschritt eine anwendungsabhängige semantische Gewichtung mittels einer normierten Relevanzfunktion zu berechnen, welche als Grundlage für die Bearbeitung bzw. das Streaming der Geodaten oder des Detailgrades des Objektes bei der Visualisierung genutzt werden kann. Durch die Zuordnung von Relevanzwerten erfolgt eine Sortierung bzw. Klassifizierung der Geoobjekte in wichtige und weniger wichtige Features (vgl. auch Zipf 2004). Hierzu werden Bewertungskriterien hinsichtlich geometrischer, semantischer, visueller und struktureller Aspekte genutzt. Diese Berechnung und die daraus abgeleitete Sortierung übernimmt eine eigene Komponente, die über den OGC WPS Relevance Sorting Service dem Grid als Dienst zur Verfügung gestellt wird.

### **WPS Terrain Generalization Service**

Multiskalige 3D-Modelle mit unterschiedlichen Level of Details (LODs) spielen eine wichtige Rolle bei Gelände- und Landschaftsmodellen. Mittels des WPS Terrain Generalization Service können Geländemodelle in verschiedenen Generalisierungsstufen erzeugt werden. In Abhängigkeit der Topographie sowie weiterer Anforderungen vereinfacht der Gelände-Generalisierungs-Dienst hochauflösende Geländemodelle. Beispielsweise genügt eine relativ niedrige Auflösung des DGM bei einer geringen Reliefdynamik, ohne typische Eigenschaften des Reliefs zu verändern. Das zugrunde liegende DGM erfährt durch die LOD-Modellierung eine Reduktion der geometrischen Komplexität (Mesh Simplification).

### 3 Fazit und Ausblick

Im Rahmen unseres Arbeitspaketes im GDI-Grid Projekt werden verschiedenen Grid Dienste bzw. gridifizierte WPS-Prozesse für die Verarbeitung von massiven Geländedaten mittels Grid Computing und herkömmlichen Geodateninfrastrukturen umgesetzt. Gleichzeitig sind Synergien zwischen der Geo-Community und der Grid-Community im Hinblick auf den Gebrauch gemeinsamer Rechen- und Speicherkapazität aber auch etablierter, standardkonformer Geodateninfrastrukturen zu nutzen. Es gilt zu beweisen, daß die Verarbeitung massiver Geodaten mit Hilfe des Grid Computing zu einem echten Performanzgewinn führt.

### 4 Literatur

- CERN (20<sup>TH</sup> APRIL, 2008): <http://gridcafe.web.cern.ch/gridcafe>
- CZAJKOWSKI, K., FERGUSON, D., FOSTER, I., FREY, J., GRAHAM, S., MAGUIRE, T., SNELLING, D. AND TUECKE, S. (2004): From Open Grid Service Infrastructure to WS-Resource Framework: Refactoring & Evolution, [http://www.ibm.com/developerworks/library/ws-resource/ogsi\\_to\\_wsrf\\_1.0.pdf](http://www.ibm.com/developerworks/library/ws-resource/ogsi_to_wsrf_1.0.pdf).
- FOSTER, I. (2002): What is the Grid? A Three Point Checklist.
- HEIER, C. AND KIEHLE, C. (2005): Geodatenverarbeitung im Internet - der OGC Web Processing Service. GIS 2005, 6: 39-43.
- SCHUT, P. (2007): OGC WPS, 2007. Open Geospatial Consortium Inc., 2007, OpenGIS® Web Processing Service, Ref. Nr. OGC 05-007r7
- SCHILLING, A., BASANOW, J., ZIPF, A. (2007): VECTOR BASED MAPPING of POLYGONS ON IRREGULAR TERRAIN MESHES FOR WEB 3D Map services. 3rd Int. Conf. on Web Inf. Systems & Technologies (WEBIST). Barcelona, Spain.
- KIEHLE, C., GREVE, K. AND HEIER, C. (2006). Standardized Geoprocessing – Taking Spatial Data Infrastructures one step further. Proceedings of the 9th AGILE 2006, Hungary.
- STOLLBERG, B., ZIPF, A. (2007): OGC Web Processing Service Interface for Web Service Orchestration - Aggregating geoprocessing services in a bomb threat scenario. Accepted for W2GIS: Web&Wireless GIS Conference, 2007.
- STOLLBERG, B. & ZIPF, A. (2008): Geoprocessing Services for Spatial Decision Support in the Domain of Housing Market Analyses - Experiences from Applying the OGC Web Processing Service Interface in Practice. 11th AGILE 2008 Confernece. Girona. Spain.
- WEISER, A., NEIS, P. AND ZIPF, A. (2006): Orchestrierung von OGC Web Diensten im Katastrophenmanagement - am Beispiel eines Emergency Route Service auf Basis der OpenLS Spezifikation. In: GIS - Zeitschrift für Geoinformatik. 09/2006. pp. 35-41.
- WEISER, A. AND ZIPF, A. (2007): Web Service Orchestration (WSO) of OGC Web Services (OWS) for Disaster Management. Joined CIG/ISPRS Conference on Geomatics for Disaster and Risk Management. 23.-25.05.2007. Toronto. Kanada.
- ZIPF, A. (2003): Zur Bestimmung von Funktionen für die personen- und kontextsensitive Bewertung der Bedeutung von Geoobjekten für Fokuskarten. Symposium für Angewandte Geographische Informationstechnologie. AGIT 2003. Salzburg.