Standards-Based Processing of Digital Elevation Models in Grid Computing Environments

Sandra Lanig[1], Stefan Kurzbach[2], Erik Pasche[2] and Alexander Zipf[1]
[1]University of Bonn, Department of Geography, Chair of Cartography
[2]Hamburg University of Technology, Department for River and Coastal Engineering

INTRODUCTION

In 2007 the European Commission passes a directive on the validation and management of flood risks. The objective of this directive is to calculate and reduce flood-conditioned risks to human health, the environment, cultural heritage and economic activity. For this purpose the directive plans a three-stage beginning. Until 2013 flood risk maps, based on flood modeling, have to generate throughout the EU. Basic requirement for these flow models in river and coastal engineering applications are high resolution *Digital Elevation Models* (DEMs). In particular the use of *Airborne Laserscanner Data* (ALS or LiDAR, *Light Detection And Ranging*) is an excellent approach to acquire high precision topographic data across a large area for generating detailed DEMs.

The geo community has established a number of standards pushed by the *Open Geospatial Consortium* (OGC). A *Spatial Data Infrastructure* (SDI) provides access to globally distributed spatial data through standard, interoperable OGC Web services in a *Service Oriented Architecture* (SOA) (Nebert, D. (Ed.) 2004). Previous standardization efforts of the OGC primarily based on discovery (*Catalogue Service for the Web*, CS-W), access (*Web Feature Service*, WFS; *Web Coverage Service*, WCS) and visualization (*Web Map Service*, WMS) of geospatial data. But the core of SDIs are analysis functions. To offer any kind of GIS processing functionality by a standardized interface, the OGC *Web Processing Service* (WPS) standard (version 1.0.0) is used. The WPS specification defines three mandatory operations: *GetCapabilities* provides service metadata and general information about the offered processes, *DescribeProcess* provides the processes' metadata including input and output parameters. *Excecute* runs an offered process.

However to process these massive LiDAR data there is a need for sophisticated data processing techniques and storage management. Linking LiDAR processing methods to conventional, already existing WPS implementations is insufficient. For this purpose WPS should be able to use *Grid Computing* systems, which provides both the computing power and the required storage capacity to achieve a high processing performance. However, some research has been accomplished in linking *Grid Computing* technology and *Spatial Data Infrastructures* (SDIs) based on OGC standards by Di et al. (2003, 2008), Baranski (2008) and Padberg (2009). Besides linking WPS to *Grid Computing* systems additionally methods to make the high quality LiDAR DEM manageable are necessary to reduce the huge number of the data points. Therefore generalization and simplification methods are used. Within the GDI-Grid project (www.gdi-grid.de) we are developing a number of Grid-enabled Web services for processing massive DEMs. For this purpose we develop adapted surface processes based on the WPS specification and the deegree framework. One of our use cases is flood modeling by hydrodynamic simulation. Here we show how massive terrain data can be processed in *Grid Computing* environments based on OGC WPS interface and synergies between *Grid Computing* and SDIs can be created by integrating OGC and Grid Computing standards in a *Service Oriented Architecture* (SOA).

USE CASE– TERRAIN PROCESSING FOR FLOOD MODELING

Increasingly drastic flood events create the requirement for detailed flood and risk maps. We use a 2D hydrodynamic numerical model in large-area time-dependent simulations to get the inundated areas over the course of a flood event. Based on these results risk maps can be produced. The spatial basis data to create a 2D flow model needs to be subjected to a special preprocessing like terrain generalization. The results can then be visualized in 3D e.g. by conversion to *Virtual Reality Modeling Language* (VRML) scenes.

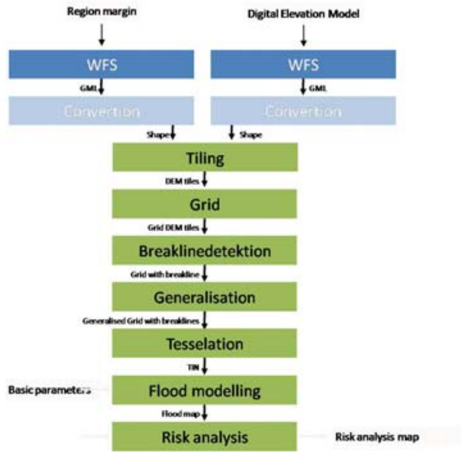

Figure 1 Terrain processing workflow for flood modeling

Figure 1 presents the terrain processing workflow for flood modeling. In the preprocessing phase massive raw laser scanning data (3D point cloud) needs to be reduced in a way that important structural terrain features (e.g. breaklines) are preserved and a *Triangulated Irregular Network* (TIN) has to be derived from the elevation points. Breakline detection is performed using image processing methods, so a raster DEM is created in an intermediate step. To minimize the huge amount of terrain data and optimize processing time, diverse generalization algorithms and different algorithms for generating *Level of Details* (LOD) are available as Grid services. The terrain discretization and simplification processes are captured in an executable workflow description.

ARCHITECTURE – GRID-ENABLED SERVICES

Within the GDI-Grid project we focus on the integration of SDI and Grid technologies. Therefore we use both technologies to develop Web services for processing high resolution massive surface data in different scales and resolutions (multi-scale spatial data, generalization) and integrate these terrain processing Web services as proof of concept in the flood simulation scenario. The system architecture of the grid-enabled WPS Terrain processing service is illustrated in Figure 2.

A *Grid* coordinates resources that are not subjected to centralized control using standard, open, general-purpose protocols and interfaces and deliver nontrivial qualities of service [Foster 2002]. Grid Computing standards like the *Web Services Resource Framework* (WSRF) enable us to make use of stateful Grid services and to submit computationally intensive tasks into the Grid. *Grid services* based on the WSRF are described by the *Web Service Description Language* (WSDL) and communicate by means of *Simple Object Access Protocol* (SOAP), both standards of the W3C. The WSRF identifies and provides a set of conventions for managing state using Web service technologies (*WS-Resource, WS-Notification*) and make Grid resources accessible within a Service Oriented Architecture.

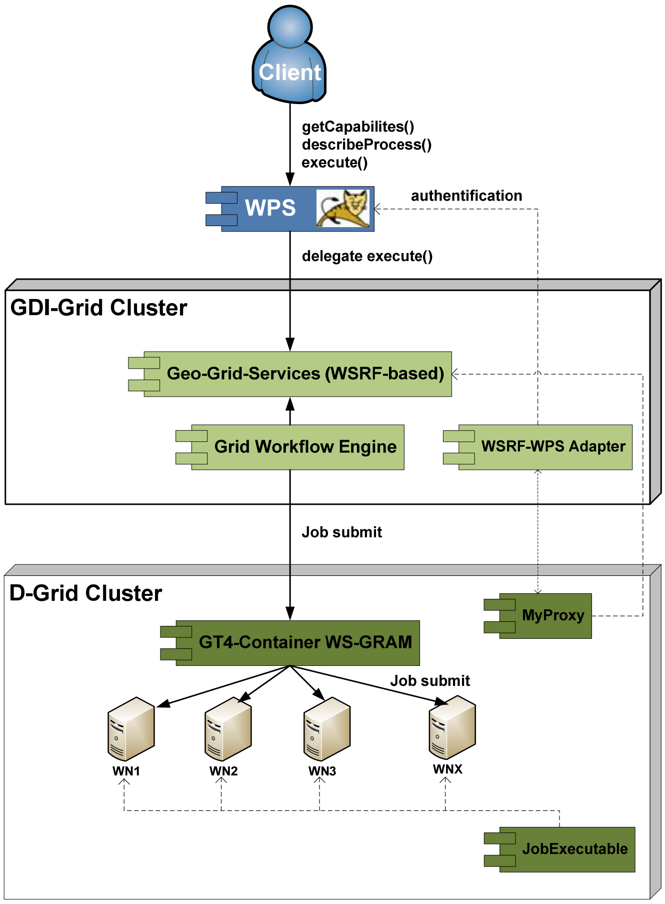

Figure 2 Terrain processing architecture

There are many differences in service discovery, description, messaging, and security methods of OGC and Grid services that lead to interoperability problems between OWS and "normal" Web services as well as Grid services. The Grid community follows the IT mainstream using SOAP and WSDL. On the other side, OGC Web services are following a restful style, but there are efforts to provide SOAP/WSDL interfaces for OWS. Additionally OWS do not support any security infrastructure like authorization and authentication methods which is a major requirement in a Grid Computing environment (*Grid Security Infrastructure*, GSI). In a first step we implement different terrain processing services based on the WPS specification. The WPS is the front-end interface to the GDI-Grid infrastructure and uses the Grid as a backend computing environment. In a second step the WPS interface is extended using GSI so the processes can be seamlessly integrated into the Grid. The services are realized as Grid services using the *Globus Toolkit 4* (GT4) framework, which are then orchestrated using a formal workflow description (*Business Process Execution Language*, BPEL) and a workflow engine capable of automatically executing the workflow in the Grid (Hobona 2007). Each service calls itself leads to a job submission to a Globus WS-GRAM, which enables us to control remote jobs on Grid-based compute resources, e.g. a computing cluster. Fleuren et al. (2008) discusses how different kinds of services from geospatial and grid domain can be integrated in a workflow. A major problem is staying OGC-compliant while at the same time supporting GSI and including legacy OWS into the workflow. We propose special *OGC Proxy Services* that access external WFS and WCS delivering the data into the Grid. Regular WS-Security mechanisms and delegation of proxy certificates ensure that the data is kept confidential. Data transfers are performed by a *Data Access and Integration Service* (OGSA-DAI) being able to efficiently gather data from a large number of different file and data base sources for processing in the Grid.

CONCLUTION AND FUTUR WORK

The need for computing power and storage capacity is steadily rising within geo-community. In particularly LiDAR data is being used to create a high-resolution 3D model of the Earth´s surface. But to process these vast LiDAR data sets means to compute million of raw data points and running computationally intensive algorithms. We have presented in this paper a possibility to make the processing of massive LiDAR data easier using standardized WPS in relation with GRID Computing. Scientists will be able to access the WPS Surface Generalization Service choosing their DEM data source and the generalization algorithm via a Web interface. Users then will be able to execute computationally intensive LiDAR data processing tasks on GRID computing resources through a standardized WPS interface.

An additional optimization is the integration of the WSRF- and WPS-based services into workflows. Fleuren (2008) discusses the problems of designing a workflow based on standard OGC Web services and WSRF-based services. For orchestrating services the use of the de facto standard BPEL (*Business Process Execution Language*) allows the construction of complex workflows. An early evaluation of this approach for OGC Web services has been provided by Weiser & Zipf 2007. After the execution the user could save and reuse the workflow description. Our future work will focus on the integration of our WPS and WSRF-based services into the geospatial workflow engine that based on the BPEL. The client can combine different LiDAR processing algorithms by creating simple XML documents.

REFERENCES

Padberg, A., Kiehle C. (2009) Geodateninfrastrukturen und Grid-Computing: Aktuelle Herausforderungen und Anwendungsbeispiele. Geoinformatik 2009, Osnabrück, Germany.

Baranski, B. (2008) Grid Computing Enabled Web Processing Service. GI-Days 2008, Münster, Germany.

Fleuren T, Müller P (2008) BPEL Workflows Combining Standard OGC Web Services and Grid-enabled OGC Web Services. Proceedings of the 34th Euromicro Conference on Software Engineering and Advanced Applications, Parma, Italy

Di, L., Chen, A., Yang, W., Liu, Y., Wei, Y., Mehrotra, P., Hu, C. and Williams, D. (2008) 'The development of a geospatial data Grid by integrating OGC Web services with Globus-based Grid technology', Concurrency and Computation: Practice and Experience, 20, (14), pp. 1617-1635.

Schut, P. (2007) OGC Web Processing Service. OGC Standard, Document Reference Number 05-007r7, Version 1.0.0.

Hobona, G., Fairbairn, D. and James, P. (2007) 'Workflow Enactment of Grid-Enabled Geospatial Web Services', Proceedings of the 6th UK e-Science All Hands Meeting. Nottingham, 10th - 13th September 2007. Workflow Enactment of Grid-Enabled Geospatial Web Services: pp. 392-399.

Sotomayour, B. & Childers, L: Globus Toolkit 4 (2006): Programming Java Services. San Francisco: Morgan Kaufmann.

Nebert D (Ed.): Developing Spatial Data Infrastructures: The SDI Cookbook Version 2.0, GSDI, p.6. [Online available from: http://gsdi.org/docs2004/Cookbook/cookbookV2.0.pdf] Accessed 10 Sept 2008.

Di, L., Chen, A., Yang, W. and Zhao, P. (2003) 'The Integration of Grid Technology with OGC Web Services (OWS) in NWGISS for NASA EOS Data', Global Grid Forum(GGF) 8 & High Performance DistributedComputing(HPDC) 12. Seattle, Washington

Foster, I. (2002): What is the Grid? A Three Point Checklist. Online unter: [Online available from: http://www.mcs.anl.gov/~itf/Articles/WhatIsTheGrid.pdf] Accessed 23 March 2009.