Hamburg University of Technology, Department of River and Coastal Engineering, University of Bonn, Department of Geography, Chair of Cartography

BENEFITS OF GRID COMPUTING FOR FLOOD MODELING IN SERVICE-ORIENTED SPATIAL DATA INFRASTRUCTURES

Stefan Kurzbach, Prof. Dr. Erik Pasche,
Sandra Lanig, Prof. Dr. Alexander Zipf

Abstract: In 2007, the European Commission has passed the "Flood Directive" (2007/60/EG) dealing with the identification of inundated areas and the creation of flood risk maps. The basis for flood modeling is provided by computationally and storage-intensive flow simulations. Digital terrain data is the starting point for generating two-dimensional flow models. When discretizing the computational mesh for hydraulic simulation, digital terrain models with a resolution of one meter or less have to be subjected to several elaborate preprocessing steps. A number of simulation and modeling tools for this purpose have already been developed as "classical" SDI applications. However, building flood and risk models for a study area covering many square kilometers is not possible using common desktop GIS. The German GDI-Grid project (SDI-Grid, www.gdi-grid.de) extends regular OGC-based SDI services with grid computing capabilities. It focuses on WPS tools and an architecture for geoprocessing in the grid. At the example of flood modeling, this article shows which benefits can be generated in spatial data infrastructures by using grid technology.

Keywords: Grid computing, flood modeling, web service, workflow, spatial data infrastructure

// VORTEILE VON GRID COMPUTING FÜR DIE ÜBERSCHWEMMUNGSMODELLIERUNG IN EINER SERVICEORIENTIERTEN GEODATENINFRASTRUKTUR

// Zusammenfassung: Die Europäische Kommission verabschiedete 2007 die Hochwasserschutzrichtlinie (2007/60/EG). Diese umfasst sowohl die Ausweisung von Überschwemmungsflächen als auch die Kartierung von Risikogebieten. Die Grundlage hierfür liefern rechen- und speicherintensive Strömungssimulationen. Digitale Geländedaten bilden den Ausgangspunkt für die Erstellung zweidimensionaler Strömungsmodelle. Bei der Diskretisierung des Berechnungsnetzes müssen Geländemodelle mit einer Auflösung von einem Meter oder weniger in mehreren Schritten aufwändig vorprozessiert werden. Zahlreiche Werkzeuge für Simulation und Modellierung wurden bereits als „klassische" GDI-Applikationen entwickelt. Für die Hochwasser- und Risikomodellierung eines viele Quadratkilometer umfassenden Untersuchungsgebiets reichen herkömmliche Desktop-GIS jedoch nicht aus. Im Rahmen des Projektes GDI-Grid (www.gdi-grid.de) erfolgt die Adaption herkömmlicher OGC-basierter GDI-Dienste für den Einsatz von Grid-Computing. Den Schwerpunkt bilden Werkzeuge und eine Architektur zur Geoprozessierung mit WPS im Grid. In diesem Beitrag wird anhand des Szenarios Hochwassermodellierung dargestellt, welcher Mehrwert in Geodateninfrastrukturen durch die Verwendung von Grid-Technologie generiert werden kann.

Schlüsselwörter: Grid Computing, Hochwassermodellierung, Webservice, Workflow, Geodateninfrastruktur

Authors

Dipl.-Inf. Stefan Kurzbach
Prof. Dr.-Ing. Erik Pasche
Hamburg University of Technology
Department of River and Coastal Engineering
Denickestraße 22
D-21073 Hamburg
E: {stefan.kurzbach, pasche}@tu-harburg.de

Dipl.-Geoinf. Sandra Lanig
Prof. Dr. Alexander Zipf
University of Bonn
Department of Geography, Institute of Cartography
Meckenheimer Allee 172
D-53115 Bonn
E: {lanig, zipf}@geographie.uni-bonn.de

1. MOTIVATION

Recent history in Europe has been shadowed by numerous flood disasters and their devastating consequences for the environment, economy, and citizens. Climatologists anticipate even more frequent and extreme precipitation events leading to extreme floods (Barredo 2007). In 2007 the European Commission acted on this issue and passed the "Flood Directive". Its scope is the evaluation and management of flood risk in all European countries. National actions are required in three steps (European Commission 2007):

1 preliminary flood risk assessment (until 2011)
2 generation of flood hazard and flood risk maps for all flood-prone river and coastal zones (until 2013)
3 preparation of flood risk management plans (until 2015).

According to the Flood Directive, flood hazard maps must display inundated areas, water depths, and flow velocities for statistical flood events of medium probability, meaning a water level or discharge that is expected to occur about every 100 years on average, as well as extreme floods, and events with lower recurrence periods. A combination of numerical simulation models and GIS has to be applied to fulfill these requirements, but the number of models to be created puts enormous pressure on the national authorities.

Simple inundation maps can be created, for instance, by extrapolation of a critical water level onto coastal areas and forelands. This can merely give a static view of the flooded areas, however, and could only be useful in the preliminary assessment step. Flood hazard maps, on the other hand, including varying water depths and flow velocities, are typically based on multi-dimensional, time-dependent flow models (also called hydraulic or hydrodynamic models) that take into account the various parameters affecting the flow situation, such as surface topography and roughness. In practice mostly one-dimensional and depth-averaged two-dimensional models are used because fully three-dimensional models have high computational requirements, and because the vertical flow component only plays a minor part in river flow (Pasche 2007).

Digital elevation models (DEM) are the main data source for flow model topography. DEMs are now readily available with a resolution of 1 meter or less. Model creation, in particular two-dimensional discretization of the flow network, is a time- and storage-consuming process and is usually carried out by consulting engineers on behalf of the national authorities (Rath 2007). A typical desktop computer is not capable of handling the data of more than a few square kilometers at a time, and it takes hours to complete a discretization process. In particular the use of high-resolution topographic data across large areas and the evaluation of the detailed simulation results creates a need for sophisticated processing techniques and storage management.

Grid computing is a technology that allows many distributed computers to collaboratively solve a single problem (Foster and Kesselman 1999). Foster has proposed a three-point checklist defining the properties of a grid. According to this list a grid "coordinates resources that are not subject to centralized control using standard, open, general-purpose protocols and interfaces to deliver nontrivial qualities of services" (Foster 2002). A grid may provide the required computational power and storage capacities for flood simulations at low cost and on demand. In this article we focus on the German D-Grid infrastructure. The application of flood modeling is investigated in the research project GDI-Grid to implement geoprocessing services within the D-Grid infrastructure using Globus Toolkit 4 and standards of the Open Geospatial Consortium. Most of the users of computing grids come from academic institutions or are associated industry partners in research projects. Not only is this due to the fact that only participants of approved projects can get access to the national grid infrastructures, but also that the access hurdle of using grid technology is very high. Solely academic institutions have the required technical know-how to overcome this barrier and to profit from the power of grid computing. Private users and small companies like consulting engineers cannot easily gain this benefit.

To overcome the problem of access to the grid and to provide the available computing resources to flood modelers we suggest the following actions:

- Supply standards-based geoprocessing services for tasks related to flood modeling,
- enable these services to utilize the power of a grid as a back-end computing and storage environment while hiding the grid's complexity, and
- integrate these flood modeling services into a service-oriented spatial data infrastructure (SDI).

2. STATE OF THE ART

We present existing practices related to service-oriented SDI and geoprocessing in grid computing environments. Integrating domain-specific services into a SDI and grid-enabling geospatial services is not limited to the field of flood modeling.

2.1 SPATIAL DATA INFRASTRUCTURES

A SDI provides access to globally distributed spatial data through standard, interoperable services in a service-oriented architecture (SOA). As a leading organization for voluntary consensus standardization the Open Geospatial Consortium (OGC) has published a number of open standards suitable for building SDIs in collaboration with the ISO/TC 211 (ISO Technical Committee 211 – Geographic Information / Geomatics). Previous efforts of the OGC have primarily been based on discovery, access, and visualization of geospatial data. However, according to Nebert (2004), a fundamental element of future SDI will be the integration of geoprocessing services, that is, processing functions that work on spatially related data. Geoprocessing services have only recently been considered in the OGC by issuing the Web Processing Service (WPS) standard. The WPS offers any processing functionality through a web-based interface via three mandatory operations. These service operations are getCapabilities for a brief service description, describeProcess returning a detailed description of selected processes, and execute for running a process (Schut 2007).

Existing geoprocessing routines, such as standard spatial algorithms (e. g. buffer, intersection, and generalization operations), can easily be wrapped as web services. Since OGC's publication of the WPS standard many reference implementations and case studies have been done. Kiehle, Greve and Heier (2007) discuss the potential of extending SDIs with geoprocessing services and state that a "generic web service architecture for providing common geoprocessing capabilities" must be established using OGC and well-known web standards.

WPS geoprocessing tasks have been implemented in several other spatial research domains e. g. for precision farming (Nash et al. 2008), simplification (Foerster and Schäffer 2007), hydrological applications (Diaz et al. 2008), biogeography (Graul and Zipf 2008), forest fire (Friis-Christensen et al. 2007), housing marketing analysis and disaster management (Stollberg and Zipf 2007, 2008), urban waste land determination and land management (Lanig et al. 2009), and terrain processing (Lanig et al. 2008, 2009). Some basic calculations like buffering are described in (Heier and Kiehle 2005). A range of processes have been implemented by the cartography research group of the University of Bonn, Germany, and have been made available at http://www.opengeoprocessing.org.

2.2 GEOPROCESSING IN GRID COMPUTING ENVIRONMENTS

A grid "coordinates resources that are not subject to centralized control using standard, open, general-purpose protocols and interfaces, and delivering nontrivial qualities of service" (Foster 2002). Grid computing infrastructures use grid middlewares for accessing and managing distributed computing and data storage resources, and to provide security mechanisms. There exist several grid middlewares. The currently most utilized and adopted middlewares are Globus Toolkit (Foster 2005), UNICORE (Uniform Interface to Computing Resources) (Streit 2009), LCG/gLite (http://www.glite.org) and dCache (Fuhrmann 2004).

In 2008, the OGC and the Open Grid Forum (OGF), an organization dedicated to the development of standards for the management of distributed computing resources as required for grid computing, have agreed to work together on harmonizing standards for geoprocessing in the grid. They have signed a memorandum of understanding concerning future collaboration (Lee and Percivall 2008). Grid-enabling geospatial processes has already been evaluated in several fields of study. Research in earth sciences strives for providing services that process sensor observations for wildfire applications as part of the GEOSS architecture (Mandl et al. 2007, Lee 2008). The CrossGrid project investigated the use of grid computing for flood forecasting (Hluchý et al. 2004).

Geoprocessing workflows and a grid processing profile for WPS are part of the OGC Web Services (OWS) Interoperability Testbed, phase 6 (OWS-6). Within the OWS-6 Baranski et al. (2009) and Schäffer and Schade (2009) deal with the chaining of geospatial processes and give guidelines for developing WPS with access to a grid computing environment. Liping Di et al. (2003, 2008), Baranski (2008), and Padberg and Kiehle (2009) give general ideas about linking grid technology and OWS. One important aspect is to overcome differences in service communication between OWS and generally SOAP- and WSDL-based grid services. Hobona, Fairbairn et al. (2007, 2009) have developed a workflow management system (Semantically-Aware Workflow Engines for Geospatial Web Service Orchestration, SAW-GEO) supporting the orchestration of grid-enabled geospatial web services.

Some research has been done on providing simulation models as geoprocessing services in a SOA. Floros and Cotronis (2006) have developed the "Service-Oriented Simulations" framework (ServOSims) aiming at composing and executing scientific simulations as stateful web services. In their model, service orchestration is based on data-centric notifications between service instances, but OGC-compliant services are not considered. Gregersen, Gijsbers and Westen (2007) designed the "Open Modeling Interface" (OpenMI) for easy definition and linking of processes in the hydrological domain. However, this approach is not based on standard web service technology, so it does not strictly fit into the SOA paradigm. The GEOSS (Global Earth Observation System of Systems) Architecture Implementation Pilot (AIP, http://www.ogcnetwork.net/AIpilot), which is part of the OGC Network, develops and deploys new process and infrastructure components for the GEOSS Common Infrastructure (GCI) and the broader GEOSS architecture based on OGC specifications.

2.3 RELATED WORK

Within the GDI-Grid project a number of WPS for the processing of digital elevation models have been developed based on the deegree framework (Fitzke et al. 2004) and have been extended for geoprocessing in a grid computing infrastructure based on the grid middleware Globus Toolkit 4 (Padberg and Kiehle 2009). Lanig et al. (2008) have shown how massive terrain data can be processed in grid computing environments based on OGC standards. We have developed a 3D Terrain Discretization Grid Service (Gaja3D) and evaluated the efficiency of this new technology at the creation of a large-scale two-dimensional flow model for the estuary of the river Elbe (Germany) from several million measured elevation points (Kurzbach and Pasche 2009). The technology presented below is based on the results and experiences of this work. We apply the WPS standard and Globus Toolkit to implement a flood modeling architecture suited for integration in a SDI that is using the German D-Grid infrastructure.

3. FLOOD MODELING BY HYDRODYNAMIC SIMULATION

A flow model represents the motion of water, e. g. in pipe networks, rivers, open channels or oceans. The common basis for all flow models is the numerical solution of the Navier-Stokes equations, a set of equations that describe the motion of fluids (Malcherek 2001). For free surface flow, as it occurs in such moving water bodies as rivers, estuaries, and oceans, two-dimensional depth-averaged models are preferred over fully three-dimensional models. This simplification results in a set of equations called shallow water equations needing to be solved by numerical methods. In some situations the flow process can be further reduced to a one-dimensional model. In order to save computation time, a combination of one- and two-dimensional (coupled) models can be applied (Schrage et al. 2009). The output of a numerical model includes time-series of variables like water depth, flow velocity, temperature, salinity, and bed load.

The numerical solution of a hydrodynamic model is based on a discretization of the surface topography and other properties affecting the flow situation, like surface roughness, vegetation, hydraulic structures (e. g. dikes, weirs, and bridges), as well as wind and waves. A two-dimensional discretization consists of a network of nodes and elements and is either a structured, regular grid or an unstructured mesh. It is usually created based on a digital elevation model of the topography and the bathymetry of a study area and has to incorporate characteristics of the terrain that are vital to the simulation (Rath 2007).

High-resolution topographic data for flood plains is nowadays gained using remote sensing methods (e. g. LiDAR). Initially, the measured points contain measurement errors, vegetation, and man-made structures. These have to be filtered prior to use. After filtering, the points are triangulated to form a continuous surface model (Triangulated Irregular Network, TIN). This TIN can, however, not directly serve as input for a hydrodynamic simulation because the number of points is much to high. In order to make high quality DEMs manageable for hydrodynamic simulation it is necessary to generalize and to simplify the underlying terrain model while preserving critical terrain features (Rath 2007).

Several algorithms are available for generating multi resolution DEMs at different levels of detail (LODs). Lanig et al. (2008, 2009) have implemented algorithms based on the research work by Garland and Heckbert (1997) as a geoprocessing service. This 3D Terrain Generalization WPS processes multi-scale DEMs in predefined LODs. The surface geometry is stored as a TIN, and the algorithm is based on an iterative generalization of edge aggregation by vertex pair contraction. The error approximation for simplification of each vertex is the sum of squared distances to the planes. This algorithm cannot be applied, however, for flow model simplification, but is rather suited for display purposes (e. g. reduction of the number of triangles for different levels of detail depending on viewer distance).

Flow models have to fulfill a number of criteria for hydrodynamic simulation. Most importantly, structural features of the terrain have to be enforced as edges in the discretization network. Other requirements may restrict the element sizes and internal angles. Structural features (e. g. breaklines or contour lines) can be derived from the DEM or can originate from external data sources. Detection of structural features is often based on a regular, rasterized version of the DEM using image processing methods (Rath 2007). This raster DEM can be interpolated from the terrain in TIN format with a resolution appropriate for the detection process.

Applying line generalization methods to the detected structural lines reduces the number of points in the resulting flow model. Based on the Douglas-Peucker algorithm Lanig et al. (2008) have implemented a 3D Line Simplification WPS. When enough structural information and a model boundary have been gathered, a constrained Delaunay triangulation is performed on the lines. Elevations in the resulting TIN are interpolated from the original DEM or from the simplified contour lines. Simulations have shown that this strategy is well-suited for flow model creation.

The geoprocessing workflow for flood modeling is depicted in Figure 1. It focuses on flow model creation. Starting with a DEM, all necessary steps for flow model creation can principally be performed automatically. Tiling the input DEM makes it possible to execute the raster creation, breakline detection and generalization tasks in parallel for different subsets of the data (denoted by three parallel arrows).

Succeeding the model creation process is the calibration of the hydrodynamic model. This means performing a possibly very large number of simulations with varying flow parameters so that the model can correctly represent one or more previously observed flow situations. Only if the calibration process has been finished successfully, the model can be used to predict the consequences of a flood event. Simulations provide the water level and flow velocity results for creation of inundation maps. The inundated areas are derived by intersection of the water levels with the original DEM. A subsequent flood risk analysis integrates vulnerability information for the flooded areas to derive a flood hazard map.

4. GRID-ENABLING SIMULATION SERVICES

Services for flow simulation and flood modeling require and produce a large volume of data. As shown above they are also based on multiple resource-intensive processing steps, which are nowadays often executed on a single computer limiting the size of flow models, blocking the computer for the time of a simulation, and cluttering the local hard drives with heaps of simulation results. Grids deliver computational power and storage capacities on demand and without the administrative effort of local computing systems. Making use of grid computing for geoprocessing and simulation tasks is thus a logical consequence. However, for adoption in an SDI the geoprocessing services should conform to the WPS standard. Many differences between OGC and grid standards concerning service discovery, description, messaging, and security methods lead to interoperability problems between OWS and grid services. 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. OGC web services, on the other hand, are following a restful service style that is in conflict with message style services like WSRF services.

In contrast to widely available and simple spatial algorithms, the majority of today's simulation models are trusted, well-tested legacy applications written in a classical scientific programming language like Fortran. Examples for two-dimensional hydraulic models include Resource Management Associates' RMA2 and RMA10 (http://www.rmanet.com), free RMA•KALYPSO developed at Hamburg University of Technology, Department of River and Coastal Engineering (http://www.tuhh.de/wb), Mike 21 by DHI (http://www.dhigroup.com), Delft3D by Deltares (http://www.deltares.nl), and others. Inputs and outputs are usually file-based, and processing is monolithic, which makes the models hard to be integrated with new technologies or to be coupled with other simulations.

Simulation services have to be grid-enabled in order to be used in the grid. Grid-enabling a part of software has become known under the term "gridification" (Lee and Percivall 2008). Aspects of gridification are making use of grid computing standards like the Web Services Resource Framework (WSRF) to develop stateful grid services, and to submit computationally intensive tasks into a computing cluster, e. g. by means of a Globus WS-Gram job submission service. Only recently there have been efforts to provide SOAP/WSDL interfaces for OWS as parts of the standards. As described in the previous section, simulation services shall be implemented as OWS, so gridification includes harmonizing or adapting the interface to the WSRF. The WPS specification has some potential to be extended with a WSRF interface thereby gaining additional capabilities. Dorka (2009) deals with the advantages of using WSRF for WPS. In the OWS-6 Grid Processing Profile engineering report (Baranski et al. 2009) we have presented our results concerning gridification of the WPS by means of the WSRF. For example, a stateful service controls and manages the submitted job and stores references to the results, which the user can later retrieve. Current developments around the WPS show that there is a need for maintaining the state of a geospatial process (Schäffer 2008). WSRF grid services provide similar functionality and the concepts to implement a stateful WPS using the WS-Resource and WS-ResourceProperties standards. Adhering to the WSRF has the additional effect that WPS developers get security "for free".

Security is a major requirement in many grid computing environments. Grid Security Infrastructure (GSI) is a specification for ensuring privacy, integrity, and delegation of privileges for communication between grid services and the user. It is used in grid middlewares like Globus Toolkit, LCG/gLite, and UNICORE. Gridification of OWS has to solve the security problem as many grid services rely on GSI. A problem is that no OWS standards support security methods like authorization and authentication in the grid. A number of possible solutions have been discussed, for instance, retrieving a stored GSI proxy certificate from a MyProxy server based on username and password credentials for a client (Padberg and Kiehle 2009, Liping Di et al. 2008). In 2007 the OGC Geo Rights Management Working Group (GeoRM, formerly GeoDRM) has issued an abstract model for rights and access management of geospatial resources (Vowles 2006). This model lacks a technical integration with W3C standards like the WS-Security specification, but a new OGC initiative strives to develop standard ways of performing web service authentication using these existing mechanisms while, at the same time, conforming to OWS standards (press release of August 4, 2009). Former security-related activities in the local German SDI North Rhine-Westphalia (GDI NRW) have resulted in the specification of the Web Authentication Service (WAS) and Web Security Service (WSS) in 2003. WAS and WSS are currently only applied in this context and have not yet been approved by the OGC.

Another aspect of gridification is that standards-based asynchronous notification mechanisms are yet missing in WPS. When a user has submitted a flow model to a flood simulation service or, likewise, started a long-running geoprocessing workflow, it is not feasible for him to wait for the results blocking his computer. For extremely large models the simulation may run for many hours if not days. The simulation service must be able to execute asynchronously and to deliver results

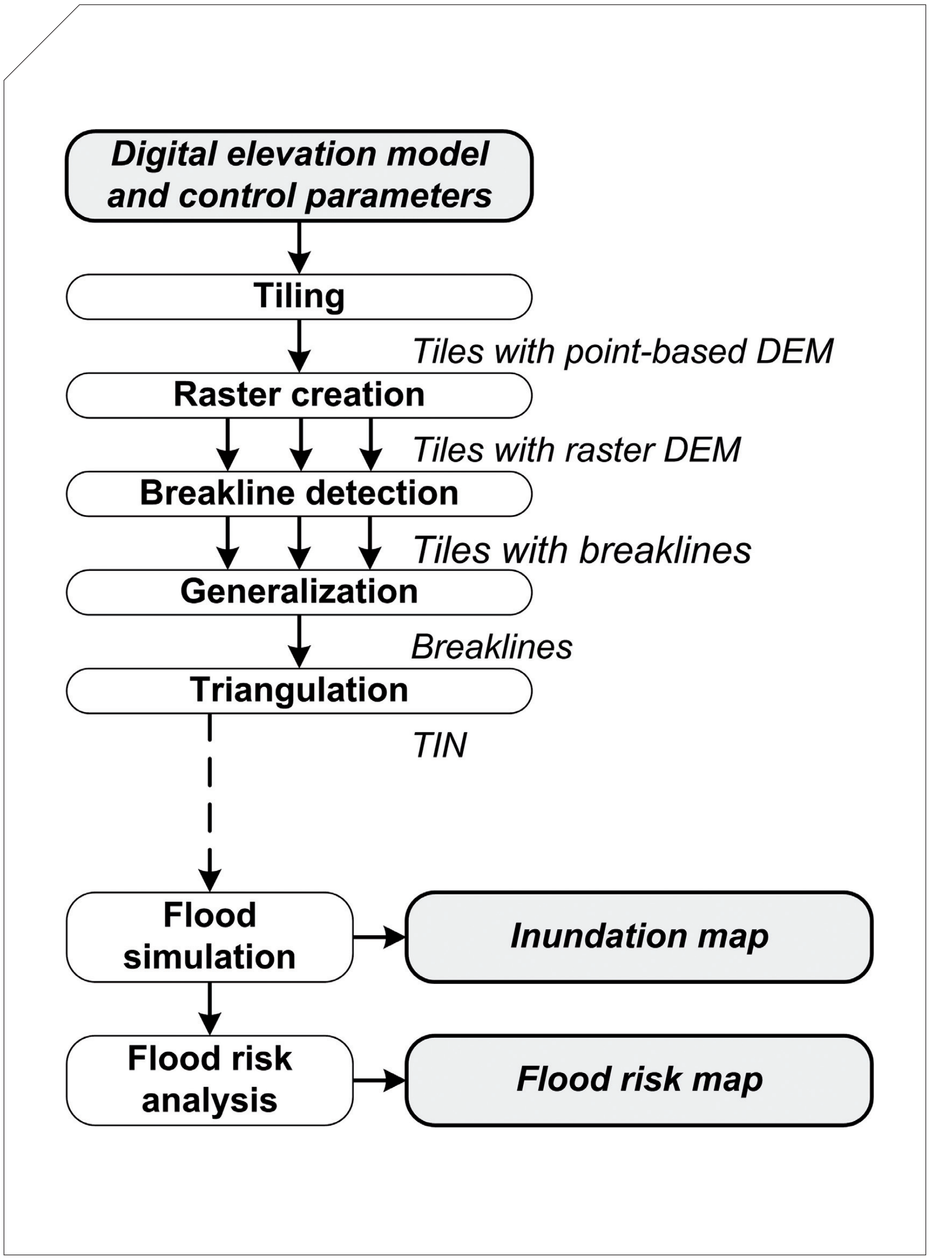

Figure 1: Geoprocessing workflow for flood modeling (focus on flow model creation)

when requested. The WPS standard supports this feature, but the client needs to poll for results. Much more convenient, and "clean" from a programmer's perspective, is an asynchronous notification using the WS-Notification standard, which is also part of the WSRF specification. For this purpose the OGC has issued the Web Notification Service standard, which is similar to WS-Notification, but uses other protocols. This makes it necessary to work on harmonizing the two standards.

5. 5 BENEFITS OF GRID COMPUTING FOR FLOOD MODELING IN A SDI

Flood simulation models are an interesting candidate for geoprocessing in an SDI. The current need for many large-scale flood simulations in Europe could be fulfilled by national flood modeling services. As the models for national rivers and potentially flooded areas are mostly non-existent, there is a need for services helping flow model creation. These could be used by engineering companies for building up the necessary models. A predefined geoprocessing workflow for flood modeling as shown in Figure 1 would further simplify the process significantly. A natural precondition is the availability of digital elevation models and other terrain data in the SDI.

By creation of geoprocessing services for legacy simulation models the functionality can be made available to a larger audience. The integration into a SDI and the specification of a standard service interface enables developers to realize an added value. There are many benefits in using grid technology for flood simulation. The most important ones from a user's point-of-view are listed below. They provide the starting point to set the requirements for our flood simulation service architecture:

- Processing on a remote machine leaving the user's computer free for other tasks,
- creation of larger flood models,
- parallel simulation of flow models,
- processing of massive terrain data,
- result management in the grid,
- keeping data confidential, securing it from unprivileged access, and
- automated execution of complete geoprocessing workflows for flood modeling.

Users as well as service developers benefit from grid technology. The existence of grid standards and their implementations makes it easier to write and to maintain better software. The existing grid middleware Globus Toolkit 4 (GT4) presents a reference implementation of the WSRF including GSI. This forms a solid base for developing standards-conforming grid services for geoprocessing and simulation as well as submission of jobs into a computing cluster. Our architecture has been designed to fulfill the mentioned requirements, but a complete practical evaluation is yet open. Nevertheless, we have quantified the efficiency of a grid service for terrain processing in (Kurzbach and Pasche 2009). The results show that a terrain dis-cretization process of the river Elbe estuary, which would take hours on a regular desktop computer, can be performed in less than 20 minutes when executed on a computing cluster. The input DEM has been partitioned into 67 tiles and separate grid jobs were run for each processing step. Input data, intermediate and final results have summed up to about 3 GB. A complete model of the river Elbe from the city of Hamburg to the estuary would be 7–8 times this size.

The grid consists of a multitude of processing and data resources that are either part of a computing cluster or single computers. A flood simulation service based on a parallel calculation core can make use of several resources for a single simulation by application of memory-parallel (OpenMP) or message-passing (MPI) communication mechanisms. Many flow models are already capable of parallel execution. However, they all lack the possibility to scale in a WSRF-based grid. Our efforts are to parallelize a flow model while respecting grid standards. We are currently developing a Flood Simulation Service that can be executed on an arbitrary number of grid nodes using standardized grid service communication and a WPS front-end for the user. Domain decomposition techniques are applied to exchange inner boundary conditions of connected model parts. Boundaries are iteratively improved to converge to a global solution. The Flood Simulation Service will be evaluated at a partitioned Elbe model that is created using the existing terrain discretization methods.

6. FLOOD SIMULATION SERVICE ARCHITECTURE

As part of our work in the GDI-Grid project, we have implemented different terrain processing services based on the WPS specification for different surface generalization functionalities. In a second step, we have extended the WPS interface using the GT4 middleware so the processes can be seamlessly integrated into the grid. The services are implemented as grid services either with GSI through MyProxy credentials or, in case of the Flood Simulation Service, as a WPS with a WSRF-conforming interface. Additional GT4 services include the 3D Line Simplification and Terrain Generalization WPS.

The geoprocessing grid services 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 (Fleuren and Müller 2008). The workflow engine contacts the Flood Simulation Service via the WSRF interface using SOAP messaging. Each grid service execution results in a job being submitted to a GT4 WS-Gram service. This enables us to control an arbitrary number of remote jobs on grid-based computing resources.

A major feature of our implementation is that no data transfers go through the workflow engine, but instead third-party transfers are initiated, and references to the results are handed over to the control of the workflow engine (see Figure 2). Data transfers are performed by a Data Access and Integration Service (OGSA-DAI) that efficiently gathers data from a large number of different file and data base sources for processing in the grid. For fast file transfers in the grid the GridFTP standard is used. WPS is the front-end interface to the GDI-Grid infrastructure and uses the grid as a back-end computing environment.

A major problem is staying OGC-compliant while, at the same time, supporting GSI and including legacy OWS into the workflow. Web Feature and Coverage Services (WFS, WCS) can now easily serve as data sources in the workflow. The OGSA-DAI server requests data from WFS and WCS, which in turn may retrieve data from a spatial data base outside the grid, and delivers the results directly to a location inside the grid. This ensures that subsequent access

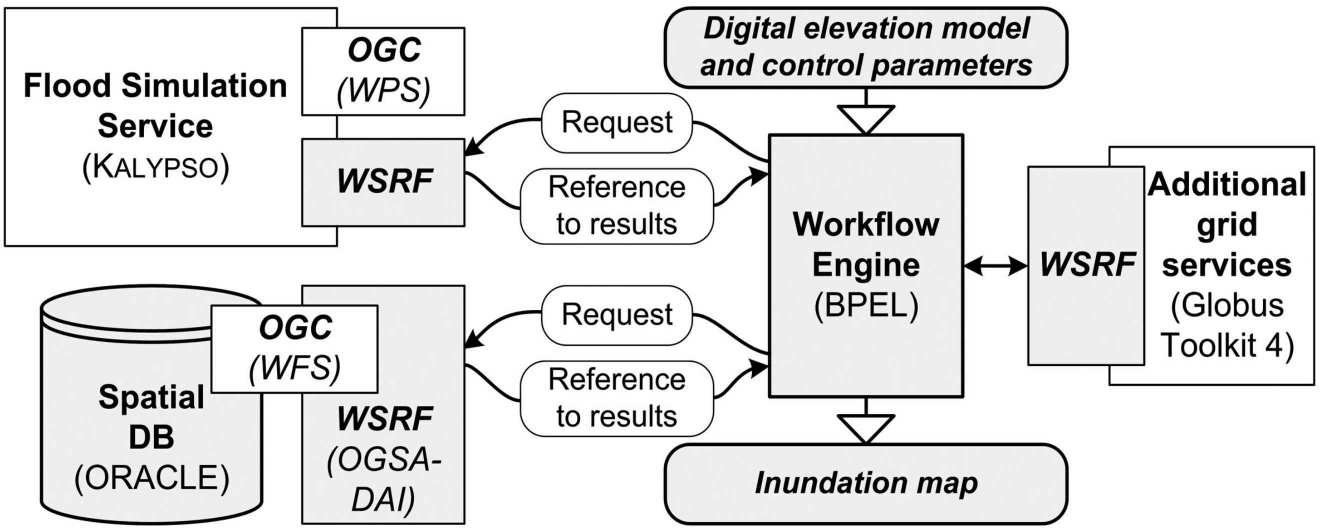

Figure 2: Distributed architecture for flood modeling service orchestration

to the data can be done efficiently based on GridFTP. Regular WS-Security mechanisms and delegation of proxy certificates to the OGSA-DAI WSRF-based service ensure that the data is kept confidential.

We have implemented a prototypical geoprocessing workflow for flow model creation in BPEL based on the workflow from Figure 1. It is executed on a Mage BPEL4Grid workflow engine with extensions for WSRF-based web services. The workflow includes retrieving data from an external WFS, processing a DEM with breakline detection and generalization as well as final TIN creation by Delaunay triangulation. Grid services have been developed with only open source software using GT4, the deegree framework, and the KALYPSO simulation platform (http://kalypso.sourceforge.net).

7. CONCLUSIONS AND FUTURE WORK

The need for computing power and storage capacity is steadily rising within the geo-community. In particular LiDAR data is being used to create high-resolution digital elevation models for flood modeling, but processing this terrain data means to work with millions of raw data points, and to run computationally intensive algorithms. In this article, we presented the possibility to enhance the processing of massive digital elevation data for flood modeling using standardized WPS and grid computing.

We also displayed how this technology can aid the creation of flow models in times of high need. The integration of grid-based geoprocessing services into a spatial data infrastructure is a logical next step. National SDIs could provide flood modeling services that help in realizing the Flood Directive, more precisely services for flow model creation, and generation of inundation and flood hazard maps. Modelers could then save time and money by using an existing grid infrastructure instead of buying expensive hardware to run their simulations. We have presented an architecture that uses WSRF-based grid services with a WPS front-end.

In future research, the management and provisioning of flood models in an SDI should be investigated. SDI and grid computing together with the appropriate tools can allow for collaborative modeling and flow model sharing. Model interfaces like the OpenMI would create the possibility to connect different flood models. If the interface is extended to create stateful WSRF-based OpenMI grid services, the shared models could then be run in the grid in a coupled fashion. This could be the future of flood modeling.

8. ACKNOWLEDGEMENTS

This work is part of the Spatial Data Infrastructure Grid project (GDI-Grid, http://www.gdi-grid.de) funded by the German Federal Ministry of Education and Research (BMBF). The main goal of the project is the efficient mining and processing of spatial data for the simulation of noise dispersion and disaster management. ◀

References

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

Baranski, B.; Shaon, A.; Woolf, A.; Kurzbach, S. (2009): OWS-6 WPS Grid Processing Profile Engineering Report. OGC 09–041r1.

Barredo, J. (2007): Major flood disasters in Europe: 1950–2005. In: Natural Hazards, Vol. 1, pp. 125–148.

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

Di, L.; Chen, A.; Yang, W.; Zhao, P. (2003): The Integration of Grid Technology with OGC Web Services (OWS): NWGISS for NASA EOS Data. GGF8 & HPDC12 2003, pp. 24–27. Science Press.

Díaz, L.; Granell, C.; Gould, M. (2008): Case Study: Geospatial Processing Services for Web-based Hydrological Applications. In: Sample, J.; Shaw, K.; Tu, S.; Abdelguerfi, M. (Edts.): Geospatial Services and Applications for the Internet, pp. 31–47. Springer Science+Business Media, New York, USA.

European Commission (2007): Directive 2007/60/EC of the European Parliament and of the Council of 23 October 2007 on the assessment and management of flood risks: Official Journal of the European Union, L 288/27.

Fleuren, T.; Müller, P. (2008): BPEL Workflows Combining Standard OGC Web Services and Grid-enabled OGC Web Services: 34th Euromicro Conference on Software Engineering and Advanced Applications. SEAA '08, Sept. 3–5, Parma, Italy, pp. 337–344. The Printing House.

Floros, E.; Cotronis, Y. (2006): ServOSims: A Service Oriented Framework for Composing and Executing Multidisciplinary Simulations. In: Kellenberger, P. (Edt.): e-Science 2006. Proceedings of the Second IEEE International Conference on e-Science and Grid Computing, 4–6 December 2006, Amsterdam, Netherlands, pp. 66–66. Wiley-IEEE, Washington DC, USA.

Foerster, T.; Schäffer, B. (2007): A Client for Distributed Geo-processing on the Web: Web and Wireless Geographical Information Systems. Proceedings of the 7th International Symposium, W2GIS 2007, Cardiff, UK, November 28–29, 2007. Lecture Notes in Computer Science, Vol. 4857, pp. 252–263. Springer, Berlin, Heidelberg.

Foster, I. (2002): What is the Grid? – a three point checklist. In: GRIDtoday, 6.

Foster, I. (2005): Globus Toolkit Version 4: Software for Service-Oriented systems. In: Jin, H.; Reed, D.; Jiang, W. (Edts.): Network and Parallel Computing. Proceedings of the IFIP International Conference, NPC 2005, Beijing, China, November 30-December 3. Lecture Notes in Computer Science, 3779/2005, pp. 2–13. Springer, Berlin, Heidelberg.

Foster, I.; Kesselman, C. (Edts.) (1999): The Grid: Blueprint for a New Computing Infrastructure.

Friis-Christensen, A.; Ostländer, N.; Lutz, M.; Bernard, L. (2007): Designing service architectures for distributed geoprocessing: challenges and future directions. In: Transactions in GIS, 6, pp. 799–818.

Fuhrmann (2004): dCache, the Overview. White paper. http://www.dcache.org/manuals/dcache-whitepaper-light.pdf, accessed August 2009.

Garland, M.; Heckbert, P. (1997): Surface simplification using quadric error metrics: Proceedings of the 24th annual conference on Computer graphics and interactive techniques, pp. 209–216.

Graul, C.; Zipf, A. (2008): Putting biogeographic research on the spatial web: Towards interoperable analysis tools for global change based on the web processing service (WPS). In: Huadong, G. (Edt.): Digital Earth Summit on Geoinformatics: Tools for Global Change Research, Vol. 1. Taylor & Francis.

Gregersen, J.; Gijsbers, P.; Westen, S. (2007): OpenMI: Open modelling interface. In: Journal of Hydroinformatics, 3, pp. 175–191.

Heier, C.; Kiehle, C. (2005): Geodatenverarbeitung im Internet-der OGC Web Processing Service. In: GIS, pp. 39–43.

Hluchý, L.; Tran, V.; Habala, O.; Šimo, B.; Gatial, E.; Astaloš, J.; Dobrucký, M. (2004): Flood Forecasting in CrossGrid Project. In: Dikaiakos, M. (Edt.): Grid Computing. Second European AcrossGrids Conference, AxGrids 2004, Nicosia, Cyprus, January 28–30, 2004. Revised Papers. Lecture Notes in Computer Science, Vol. 3165, pp. 51–60. Springer, Berlin, Heidelberg.

Hobona, G.; Fairbairn, D.; Hide, H.; James, P. (2009 [Im Druck]): Orchestration of Grid-Enabled Geospatial Web Services in Geoscientific Workflows. In: IEEE Transactions on Automation Science and Engineering, 99.

Hobona, G.; Fairbairn, D.; James, P. (2007): Workflow Enactment of Grid-Enabled Geospatial Web Services. Proceedings of the UK e-Science All Hands Meeting.

Kiehle, C.; Greve, K.; Heier, C. (2007): Requirements for Next Generation Spatial Data Infrastructures-Standardized Web Based Geoprocessing and Web Service Orchestration. In: Transactions in GIS, 6, pp. 819–834.

Kurzbach, S.; Braune, S.; Grimm, C.; Pasche, E. (2009): Hochwasser-Modellierung im Geodateninfrastruktur-Grid. In: Strobl, J.; Blaschke, T.; Griesebner, G. (Edts.): Angewandte Geoinformatik 2009: Beiträge zum 21. AGIT-Symposium Salzburg. July 7–10, Salzburg, Austria. Wichmann Verlag.

Kurzbach, S.; Pasche, E. (2009): A 3d Terrain Discretization Grid Service for Hydrodynamic Modeling: Proceedings of the 8th International Conference on Hydroinformatics (HEIC), Concepción, Chile.

Lanig, S.; Klärle, M.; Zipf, A. (2009): Offenes, nachhaltiges Landmanagement mittels standardisierter Web Processing Services (WPS). In: Strobl, J.; Blaschke, T.; Griesebner, G. (Edts.): Angewandte Geoinformatik 2009: Beiträge zum 21. AGIT-Symposium Salzburg. July 7–10, Salzburg, Austria. Wichmann Verlag.

Lanig, S.; Schilling, A.; Stollberg, B.; Zipf, A. (2008): Towards Standards-Based Processing of Digital Elevation Models for Grid Computing through Web Processing Service (WPS). Work ▷

shop on Information Systems and Information Technologies (ISIT 2008). In: Gervasi, O.; Laganà, A.; Murgante, B.; Taniar, D.; Mun, Y.; Gavrilova, M. (Edts.): Proceedings of the International Conference on Computational Science and its Applications. Perugia, Italy, June 30 – July 03. Lecture Notes in Computer Science, Vol. 5073, pp. 191–203. Springer, Berlin, Heidelberg.

Lanig, S.; Zipf, A. (2009): Towards generalization processes of LiDAR data based on GRID and OGC Web Processing Services (WPS): Geoinformatik 2009, Osnabrück, Germany.

Lee, C. (2008): Grid Technologies for GEOSS: The User and the GEOSS Architecture XXV. Perspectives on GEOSS Architecture: Principles and Implementation. Workshop on OGC TC Plenary, December 3–4th 2008, Palacio de Congresos, Valencia, Spain.

Lee, C.; Percivall, G. (2008): Standards-Based Computing Capabilities for Distributed Geospatial Applications. In: Computer, 11, pp. 50–57.

Malcherek, A. (2001): Hydromechanik der Fließgewässer. Bericht Nr. 61. Habilitation, Hannover.

Mandl, D.; Sohlberg, R.; Justice, C.; Ungar, S.; Ames, T.; Frye, S.; Chien, S.; Cappelaere, P.; Tran, D. (2007): Experiments with user centric GEOSS architectures. In: Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International, pp. 297–300.

Nash, E.; Korduan, P.; Bill, R. (2007): Optimising data flows in precision agriculture using open geospatial web services. In: Precision Agriculture, pp. 753–759.

Nebert (2004): Developing Spatial Data Infrastructures: The SDI Cookbook, Version 2.0.

Padberg, A.; Kiehle, C. (2009): Geodateninfrastrukturen & Grid-Computing: Aktuelle Herausforderungen und Anwendungsbeispiele. In: Reinhardt, W.; Krüger, A.; Ehlers, M. (Edts.): Geoinformatik 2009, pp. 129–138.

Pasche, E. (2007): Flood Modelling in Urban Rivers – the State-of-the-Art and Where to Go. In: Ashley, R.; Garvin, S.; Pasche, E.; Vassilopoulos, A.; Zevenbergen, C. (Edts.): Advances in Urban Flood Management, pp. 59–89. Taylor & Francis, Leiden, The Netherlands.

Rath, S. (2007): Model discretization in 2D hydroinformatics based on high resolution remote sensing data and the feasibility of automated model parameterisation. Dissertation. 1st ed. Hamburger Wasserbau-Schriften. Hamburg University of Technology, Hamburg, Germany.

Schäffer, B.; Schade, S. (2009): OWS-6 Geoprocessing Workflow Architecture Engineering Report. 07–143r1.

Schrage, N.; Antanaskovic, D.; Jung, T.; Pasche, E. (2009): KALYPSO – An Open Source Software Tool for Flood Studies in Rivers: Proceedings of the 8th International Conference on Hydroinformatics (HEIC), Concepción, Chile.

Schut, P.; Whiteside, A. (2007): OGC Web Processing Service (WPS) Specification. OGC 05–007r7.

Stollberg, B.; Zipf, A. (2007): OGC Web Processing Service Interface for Web Service Orchestration – Aggregating Geo-processing Services in a Bomb Threat Scenario: Web and Wireless Geographical Information Systems. Proceedings of the 7th International Symposium, W2GIS 2007, Cardiff, UK, November 28–29, 2007. Lecture Notes in Computer Science, Vol. 4857, pp. 239–251. Springer, Berlin, Heidelberg.

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. In: Friis-Christensen, A.; Pundt, H.; Compte, I. (Edts.): Proceedings of 11th AGILE International Conference on Geographic Information Science. Taking Geoinformation Science One Step Further. Girona, Spain, May 5–8.

Streit, A. (2009): UNICORE: Getting to the heart of Grid technologies. In: eStrategies – Projects, 9th edition, March 2009, pp. 8–9.

Vowles, G.; Herring, J.; Goodwin, C.; Scardina, M.; Matheus, A.; Opincaru, C.; Debeugny, W. (2006): OGC Geospatial Digital Rights Management Reference Model (GeoDRM RM). OGC 06–004r3.