Grid-based optimization in groundwater management

O. Arndt[1], U. Junghans[1], S. Kaden[1], P. Schätzl[1] and F. Thilo[2]

[1] DHI-WASY GmbH
Member of DHI group
Waltersdorfer Str. 105, D-12526 Berlin, Germany

[2] Information Systems Institute, University of Siegen,
Hölderlinstr. 3, D-57068 Siegen, Germany

Abstract
A common technique for groundwater management modeling is the Finite-Element method. Optimizing these complex models often implies many hundreds of model simulations and due to the time consuming simulation runs this leads to an impracticable runtime of the optimization process even on modern computation resources. In this paper an approach is presented to use a grid infrastructure to calculate the single simulation task in parallel. Initial results are presented, which demonstrate the feasibility of the grid-based optimization approach.

Keywords simulation, groundwater, grid, optimization, Finite-Element method

Introduction
In the field of groundwater management as well as many other engineering domains Finite-Element methods are widely used, e.g. to model environmental systems, especially subsurface flow and transport processes in groundwater. Increasingly large areas and complicated hydrogeological structures have to be simulated and are often represented by complex 3-dimensional Finite-Element model (FEM) meshes. Soaring, this leads to increasing computational demands and therefore to time-consuming simulations. At the same time the performance increase of traditional single CPU systems is beginning to decelerate. Therefore, new techniques have to be used to solve complex groundwater modeling tasks on parallel computing resources.
To speed up groundwater simulation two principle approaches can be used: applying fine-grained parallelism in the core of the simulation software itself or to execute several simulations simultaneously, thus utilizing parallelism on a more coarse-grained level. The former can be used to potentially speed up any simulation, but requires changes to the solver code. Also the scalability is limited due to synchronization overhead at the decomposition boundaries. The latter can be applied, whenever several simulations would have to be run consecutively in a traditional, sequential setup and the input of one simulation does not depend on the result of another. This paper will concentrate on the second approach for parallelization.
In the past, mainly parallel computing in groundwater modeling was rare. If it took place, networks of workstations or compute clusters were the architecture of choice. The main disadvantage of these cluster-based solutions is that in general the scalability is limited by the size of the cluster and not by the problem size, i.e. the full potential of parallelism cannot be exploited. Also the installation and maintenance of clusters can become very costly. Thus, an appropriate infrastructure is required to allocate resources on demand. Consequently there is only a limited availability of

such clusters for practical engineering tasks. Grid technology (Foster and Kesselman 2004) is a promising approach to implement a distributed resource infrastructure which is not limited to single clusters, sites or administrative boundaries.
In this paper, an approach is presented to provide the simulation core of the FEFLOW® (Diersch 2007) groundwater modeling and simulation application and the OpTiX suite (Barth et al 2000) as grid-services in a Globus® Toolkit (globus.org 2008) grid environment to enable model-based optimization.

Grid-based architecture

The latest Globus Toolkit version 4 (GT4) is based on the Service Oriented Architecture (SOA) (Singh and Huhns 2005) software engineering concept. To design GT4 compliant services, the service interfaces are described using the Web Service Description Language (WSDL) which is a XML-definition to describe a service including types, messages, and bindings developed from the World Wide Web consortium (W3C.org). Based on the SOA concept several services like groundwater simulation, distributed optimization and sensitivity analysis were implemented and deployed to resources in the grid.
To solve complex groundwater modeling tasks, these services are orchestrated to perform the necessary workflow and solve the required simulations on distributed grid simulation resources. Different services can be provided by particular experts, e.g. an optimization expert will offer several optimization algorithms as a service, while a resource provider offers low-level computation services. Even software licensing may be offered as service to avoid the inflexibility of having each potential user to purchase the required licenses a-priori. Linking the service offers of different providers promises the potential to solve even complex and otherwise extremely time consuming groundwater engineering tasks in an adequate time.
Based on the Globus grid middleware platform, several services have been implemented which can be used to support common workflows in groundwater engineering. The main services are:

a) A groundwater flow, heat and mass simulation service which is based on the simulation core of the FEFLOW® groundwater modeling and simulation application.
b) An optimization service which utilizes the OpTiX suite as a backend and offers a variety of optimization algorithms which are particularly suited for parallel optimization of constrained simulation-based optimization problems. The optimization algorithms include Distributed Polytope Search (Barth et al 2000), Asynchronous Parallel Pattern Search (Hough, Kolda and Torczon 2001), Parallel Scatter Search (Marti, Laguna and Glover 2006), and implementations of Evolution Strategies (Ahn 2006) and Particle Swarm Optimization (Kennedy, Eberhart and Shi 2001) which can evaluate several solution candidates in parallel.
c) A service to generate different model parameterizations to estimate the effect of uncertain parameters like conductivity and storability and
d) a service which is able to evaluate the results of the multiple simulations, for example to calculate the deviation of the parameter variations.

Finally a client component was integrated into the FEFLOW user interface to enable potential customers like groundwater experts to access the grid environment from their familiar working environment.
Each service can be offered by independent providers at different locations, e.g. a resource provider will offer the core simulation service while an optimization expert offers the optimization service. The optimization workflow can be composed by the

end user using the preferred providers. To enable secure resource usage, Globus suppose the user to have a valid X.509 certificate which is accepted by the resource provider, e.g. due to an out-of-band agreement.

Practical application

During the groundwater modeling workflow, series of model applications are common at different lifecycles of the model. These include the calibration of model parameters, sensitivity analysis of parameters which have to be estimated, and technical model based optimization. A typical example for the last application field is to minimize operational costs of pumping stations while still satisfying constraints like a minimum allowed depth to water table or to meet certain groundwater quality thresholds. Each of the tasks above can lead to a large number of model simulation runs, which require no interconnects apart from modifying the parameters and returning the simulation results. Hence, these tasks are perfectly suited to be executed in parallel.

Another promising approach of using grid technology takes place during the operational usage of the groundwater model for control and optimization tasks. In many areas man-made modifications caused dramatic environmental consequences, for example a change in ground water table caused by water gates or mining activities. To act against a groundwater rise, in many areas pumping stations must operate to protect settlement areas and parks or to retain the course of rivers. In general such pumping stations are built as the need arises, and thus they are often located and operated in a suboptimal manner. Driven by rising energy costs it becomes more and more imperative to reduce operational costs, e.g. by implementing optimal extraction strategies and by optimizing location placement in case of pumping station substitution. To do so, FEM-based optimization is one of the most promising approaches. But even for relatively small optimization tasks, the large number of required simulations prohibits the use of sequential and often also cluster-based optimization of groundwater models. Using the grid services it was possible to demonstrate the feasibility of the grid-based optimization approach by reducing the operational dewatering costs using an optimized pumping station placement.

The working model is based on an existing groundwater model representing the quaternary aquifer between river Rhine and river Niers in the Lower Left Rhine Area. Wide parts of this aquifer are influenced by soil subsidence which is caused by coal mining activities. During decades of coal mining surface and subsurface dewatering structures were build up as need arises. In near future mining activities in this area will be given up so stabile underground structure is going to allow a long term planning of dewatering system. A test area was chosen to proof the general usability of the selected optimizing algorithm in the field of location optimization. The test area is very heterogeneous in land use and hydrogeology which has to be respected during the optimization process. Objective of the optimization task is the overall minimization of pumped water in the test area. Pumping station types are defined by allocating delivery rates. These pumping station types form the decision variables along with the real world coordinates of the locations. As constraint a minimum allowed depth to water table depending on land use was used. This constraint was mapped as a water table which was generated using a digital terrain model (DTM) and a land use map. It is stored in the FEFLOW ground water definition file as a reference distribution. The constraint must be satisfied in the whole test area but pumping station placement may be restricted. So areas with low conductivities and territories which can not be used as pumping station location due to different reasons were defined as “no go zones“ stored as a reference distribution as well.

A graphical user interface was implemented as a FEFLOW module to enable FEFLOW users to define a location optimization task definition, even if they are not used to work with a command line shell. The FEFLOW grid graphical user interface can be accessed from the FEFLOW client application. It offers user-friendly dialogues for problem definition and optimization job control, thus allowing the modeler to use the grid features from within the environment he is used to.
The aforementioned grid services were deployed to two sites at WASY and the University of Siegen, respectively to solve this optimization problem. Parallel Scatter Search (PSS) was selected as the optimization algorithm for its generally high robustness, thus making it a good candidate when tackling a new problem. The problem was configured via the GUI with a maximum of 50 extraction wells, resulting in a total of 150 decision variables (2D coordinates + extraction rate per pumping station). The number of installed CPUs at the sites is 34, but because the optimization job had to compete with other computations on the same resources, only half of them could be used on average. At the time of writing this paper, the calculations are still ongoing: 3300 distinct combinations of well locations and extractions rates have been simulated so far at an average simulation time per candidate of about 30 minutes, resulting in an aggregate sequential computation time of approximately 1650 hours. Because of the use of parallel resources, only 80 hours of real time have been spent so far, though. Based on past experience with the optimization algorithm it can be roughly estimated that no more than 20% of the optimization is finished at this point, giving us a remaining time of 400 more hours. As can be seen from these numbers, a purely sequential computation would be infeasible because of the excessive time requirements. The time could be further reduced if more resources were available: up to about 300 simultaneously running FEFLOW instances could probably be used effectively for the given problem, reducing the total runtime by another factor of 10. Because the optimization is still in its early stages, no optimization result can be presented here. The intermediate results indicate that the initial random solution could already be significantly improved upon (average violation of depth to water table constraint was reduced from 1.5m to 1.2m) and further progress is to be expected. The approach and its technical implementation have already proven their practicability.

Conclusions and future work

In this paper, an approach was presented to apply the grid technology to FEM-based optimization, and therefore to access distributed resources. A grid infrastructure was set up and the FEFLOW core simulation system and different optimization algorithms were integrated as grid services. It was possible to define a workflow of FEM-based optimization in a grid environment and to demonstrate the feasibility of the approach. It was also demonstrated, that sequential computation is not practicable to solve such FEM-based optimization tasks within a reasonable time.
Nevertheless there is still some work to do. Currently, the ongoing work is the optimization of the placement of pumping stations as described in the practical application section, which is a completely new domain within groundwater modeling. So far it was possible to demonstrate that the technical implementation is practicable, but it must also be validated that the economic (e.g. energy cost of pumping stations) and environmental (e.g. allowed depth of water table) results can be improved with respect to manual decisions.
The next step will be to analyze the results, to improve the optimization settings, and thus to demonstrate the improvement which can be obtained by using FEM-based optimization. That includes comparative studies of different algorithms and models as

well as scalability analysis using additional resources. It will also answer the question, whether the algorithms currently implemented in the OPTIX suite are qualified to solve the optimization task of e.g. well placement or if alternative algorithms are required. It will also be necessary to educe general recommendations for defining and performing FEM-based optimizations from the initial experience of solving such tasks.
Further results of the ongoing work will be presented during the presentation.

Acknowledgements
This research is supported within German Grid Initiative (D-Grid) by the German Federal Ministry of Education and Research (BMBF) under the contract 01AK806E and within BEinGRID Business Experiments in GRID by the European Commission under the contract number 034702.

References

Ahn, C.W.: “Advances in Evolutionary Algorithms: Theory, Design and Practice”, Springer, 2006.

Barth, T., Freisleben, B., Grauer, M., and Thilo, F.: “Distributed Solution of Simulation-Based Optimization Problems on Networks of Workstations”, Journal on Computer Science and Systems, pp. 94-105, 2000.

H.-J. Diersch: “FEFLOW 5.3 - Finite Element Subsurface Flow & Transportation System”, User’s Manual, WASY GmbH, Berlin, 2007.

Foster, I., and Kesselman, C. (Eds.): “The Grid 2: Blueprint for a New Computing Infrastructure”, Morgan Kaufmann, 2004.

globus.org: The Globus Alliance: http://www.globus.org (visited 2008-03-12)

Hough, P., Kolda, T. G., and Torczon, V.: “Asynchronous Parallel Pattern Search for Nonlinear Optimization”, SIAM Journal of Scientific Computing 23(1), 134-156 (2001).

Kennedy, J., Eberhart, R., and Shi, Y.: “Swarm Intelligence”, Morgan Kaufmann Publishers, 2001

Marti, R., Laguna, M., and Glover, F.: “Principles of Scatter Search”, European Journal of Operational Research, Vol. 169, Issue 2, pp. 359-372, 2006.

Singh, M., P.; Huhns, M., N.: “Service-Oriented Computing Semantics”, Processes, Agents, Wiley 2005.

W3C.org: World Wide Web consortium: http://www.w3.org (visited 2008-03-13)