University of Tübingen
Center for Applied Geosciences
GeoSystemsResearch

APPLICATION OF HIGH PERFORMANCE COMPUTING TECHNIQUES (PARALLEL PROCESSING) TO THE MODELING OF COMPLEX COUPLED GEO-PROCESSES USING A FINITE ELEMENT APPROACH (GEOSYS/ROCKFLOW)

Dany Kemmler and Olaf Kolditz
(Center for Applied Geoscience, University of Tuebingen, Germany)

Panagiotis Adamidis and Rolf Rabenseifner
(High Performance Computing Center, University of Stuttgart, Germany)

GeoSys – Preprint
[2004-52]

Tübingen, Jan. 2005

University of Tübingen
Center for Applied Geosciences
GeoSystemsResearch

APPLICATION OF HIGH PERFORMANCE COMPUTING TECHNIQUES (PARALLEL PROCESSING) TO THE MODELING OF COMPLEX COUPLED GEO-PROCESSES USING A FINITE ELEMENT APPROACH (GEOSYS/ROCKFLOW)

Dany Kemmler and Olaf Kolditz

(Center for Applied Geoscience, University of Tuebingen, Germany)

Panagiotis Adamidis and Rolf Rabenseifner

(High Performance Computing Center, University of Stuttgart, Germany)

GeoSys – Preprint
[2004-52]

Tübingen, March 2005

APPLICATION OF HIGH PERFORMANCE COMPUTING TECHNIQUES (PARALLEL PROCESSING) TO THE MODELING OF COMPLEX COUPLED GEO-PROCESSES USING A FINITE ELEMENT APPROACH (GEOSYS/ROCKFLOW)

Dany Kemmler and Olaf Kolditz
(Center for Applied Geoscience, University of Tuebingen, Germany)
Panagiotis Adamidis and Rolf Rabenseifner
(High Performance Computing Center, University of Stuttgart, Germany)

Abstract: *This paper describes a finite element geo-process modeling software, which is able of solving multiphysics problems in the area of geo science. First results of a water resources management model for the Jordan Valley area are presented. These kind of problems are very demanding in terms of CPU time and memory space, which are typically not available on a single processor. Therefore a parallelization strategy based on domain decomposition is proposed, so that the usage of multiple processors of a parallel computing system is possible.*

Introduction

Modeling complex coupled transient and dynamic processes found in geo-systems requires increasingly high resolution models for a more precise and qualified validation and evaluation. The more exact the investigation of the given problem can be accomplished, the better the utilization, management and planning of the given geo-resources can be undertaken.

GeoSys/RockFlow (GS/RF) [1] is a finite element geo-process modeling software tool finding increased application in modeling and simulation in fields such as water resource management, geotechnics, design of geo-engineered barriers, exploitation of geothermal energy, soil and groundwater contaminant transport and remediation strategies. Its current design is based upon a sequential approach.

Here we present a new project which started summer 2004 and is funded by the German Research Foundation (DFG) and the State of Baden-Wuerttemberg (BW). Its goal is to bundle the knowledge of the experts at the High Performance Computing Center (HLRS) in Stuttgart and the scientists and researches working at the Center for Applied Geology, (ZAG) at the University of Tuebingen. This project is truly interdisciplinary and multinational as specialists in mathematics, computer science, hydrology, geo science, chemistry, just to name a few, are continuing to develop the already existing serial software and to transfer GS/RF into a highly performing massive-parallel system.

GeoSys/RockFlow Model Features

Processes. The finite element simulator GS/RF covers a wide range of physical and chemical processes relevant to environmental hydrosystems. The processes can be grouped in 4 different categories as summarized below:

1. **Hydrological Processes:**
   - Groundwater flow in confined and unconfined aquifers
   - Multi phase flow
   - Fracture flow, dual porosity
   - Density dependent flow (thermal, tracer)
   - River flow (based on averaged 1-D Saint-Venant equations)

2. **Thermal Processes:**
    - Heat transport with density changes
    - Non isothermal multiphase flow with phase changes
3. **Chemical Processes:**
    - Multi-component transport with density changes
    - Sorption models
    - Reactive Transport (i.e Freundlich Isotherm)
    - Chemical reactions via coupling to PHREEQC2
4. **Mechanical Processes:**
    - Poro elasticity
    - Thermo elasticity
    - Elasto plasticity (hardening)

Pre- and Post-Processing. Models can be created and run using a graphical user interface (Windows application). Built-in mesh generators are: gmsh, PrisGen and TetMesh. Meshing of complex structures can be done using gOcad or the pre-processor GINA developed by Federal Institute for Geosciences and Natural Resources (BGR). Hybrid meshing is possible. ArcGIS shape files can be read and converted to polylines which are then used to create meshes or assign boundary or initial conditions etc. An interface to Gstat allows for the generation of three dimensional heterogeneous conductivity fields. Two dimensional contour plots and time-value graphs are displayed during the simulation in the GUI or Tecplot output files are written directly or created with the post-processor RF2TP also developed by BGR.

Software Architecture. Despite its long history dating back to 1985 the code is programmed according to recent programming principles. The program was rewritten in ANSI-C in 1996 to enable the use of dynamic data structures and object-oriented programming. In 2003/2004 the code was again re-organized to benefit even more from object-orientation and to allow an easier switching between process couplings. Most recent changes are: use of C++, organization of RockFlow into GeoSys: GEOLib, MSHLib, FEMLib and the creation and encapsulation of process-oriented objects (PCS). These changes provide a solid basis for further program development within a growing research team.

Model of Jordan Valley. As an example for the necessity of high performance computing in environmental science we present first results of a water resources management model for the Jordan Valley area [2]. Fig. 1 shows the structural model of the investigation area which consists of 84 geometric entities (volumes) based on geological considerations. Geometric complexity is one of the features of those environmental systems. This requires spatial high-resolution discretisations in the order of several million grid points. Second is the complexity of physico-chemical processes resulting in the solution of coupled non-linear multi-field problems with many degrees of freedom in terms of unknown field functions.

The hydraulic system is controlled by recharge and discharge conditions to or from the model area as well as by discharges from several springs. Fig. 2 shows the distribution of the steady state hydraulic head in the model domain. Highest values occur in the northern recharge area. Lowest values are in the western spring field.

Geothermal basic processes are illustrated in Fig. 3. There is a permanent heat flow from the base to the system. Through the North groundwater is entering the upper aquifer and through the East the lower aquifer. This groundwater has to cool the whole system; otherwise the temperature will increase permanently. Fig. 3 shows a first long-term simulation (30.000 years) of the thermal system based on the hydraulic model presented above. The simulation shows a permanent increase of temperatures. The groundwater entering the system is not equilibrating the base heat flux. This indicates to possible defects in the current model. First, the outside groundwater recharge to the domain is underestimated. Second, the base heat flux is overestimated. This means, involving thermal data to the simulation, the hydraulic model can be improved.

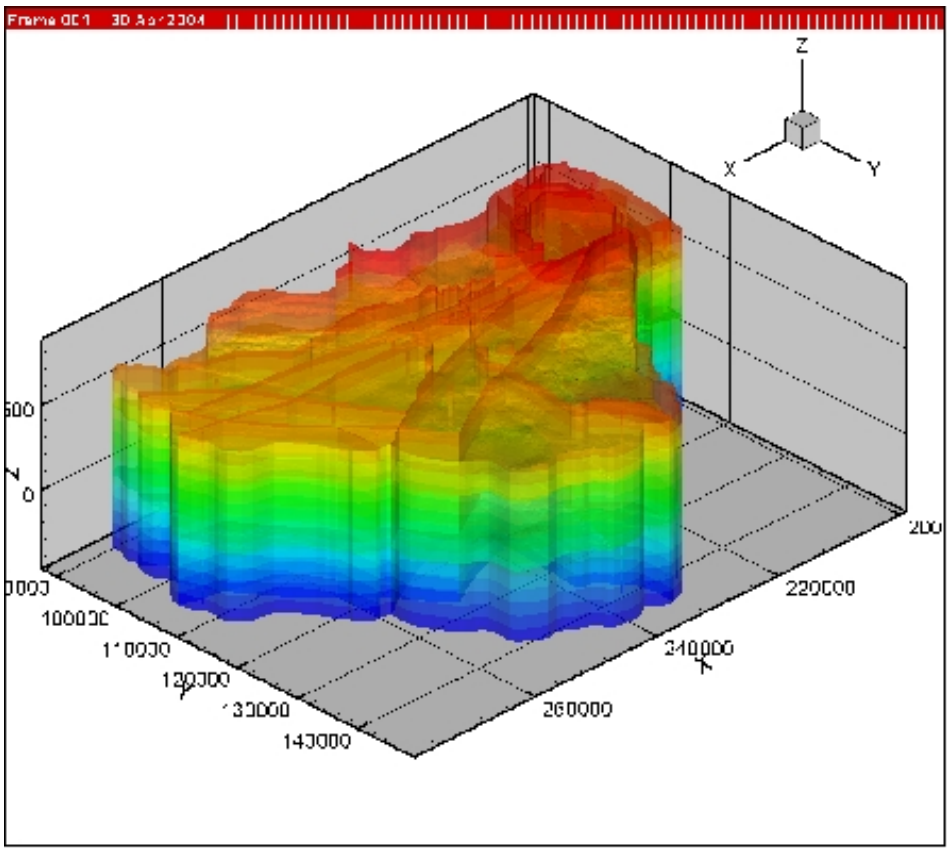

Figure 1: Structural model of the investigation area of Jordan Valley

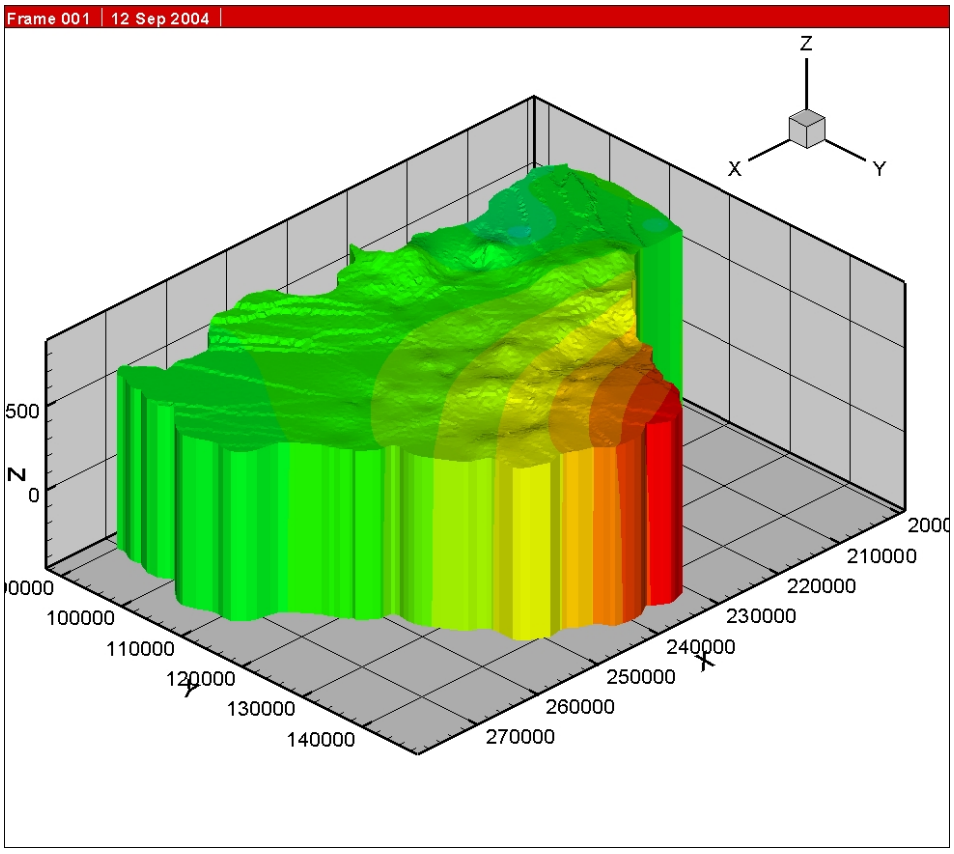

Figure 2: Distribution of the steady state hydraulic head in the model domain

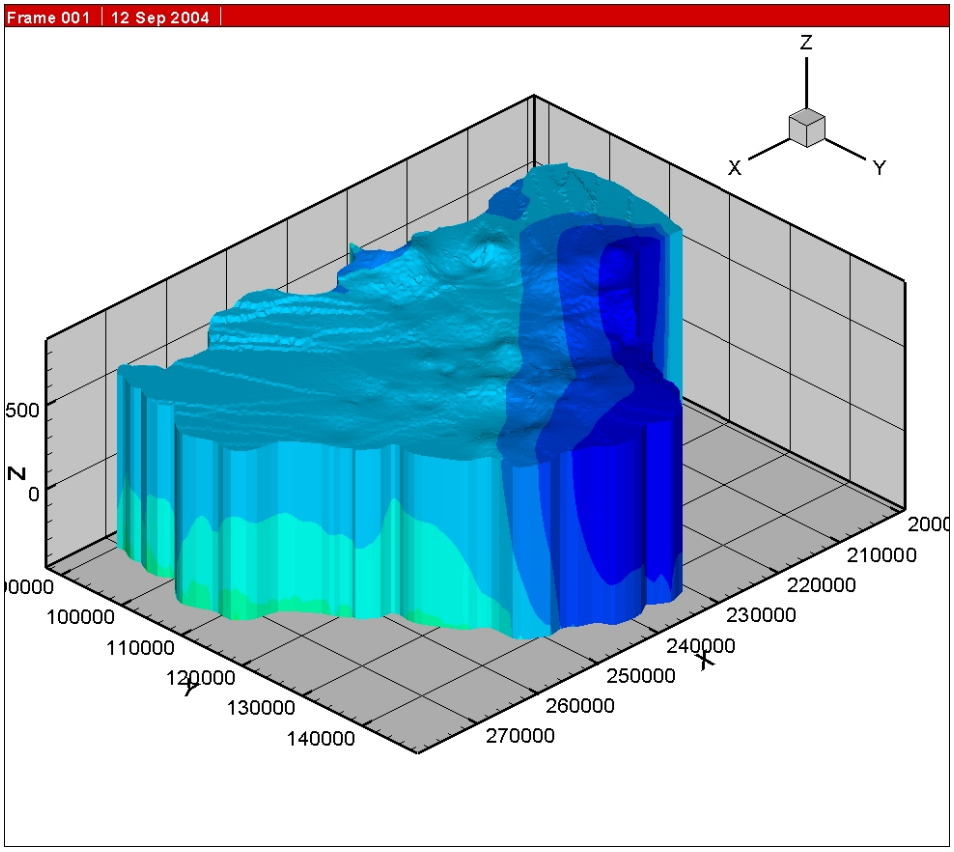

Figure 3: Long-term simulation of the thermal system based on the according hydraulic model

Parallelization Strategy

Todays modern single processor machines and desktop computers are improving in processor speed and performance as well as constantly increasing storage media. However, they are not powerful enough to process large scale models necessary for the realistic simulation of in situ highly complex geo-science systems.

Through the combination of the wide field of geo-science and high performance computation supported through effective algorithms found in computer science a complete new level in simulation and modeling can be attained. The final goal is to compute models consisting of more than 10 millions of elements with a higher grade of specialization.

The target parallel platform for the work presented in this paper is a parallel architecture with distributed memory, in particular cluster of SMP's. Such clusters consist of several nodes, where each node has a number of processors (eg. 2 or 8) with shared memory. In order to run an application on such a computational environment the work and the data have to be distributed among the processors. The decision that has to be made is which work will be done on what data. In case each process is performing a different task on different data we are talking about a MPMD (Multiple Program Multiple Data) programing model. If on the other hand the same program is running on every processor but with different data, this corresponds to a SPMD (Single Program Multiple Data) programing model. In this work we decided to use the SPMD paradigm.

Domain decomposition techniques are a very suitable way of decomposing data and work in numerical analysis. Using these techniques we can divide the original computational domain into a number of subdomains, equal to the number of available computing units, and distribute these subdomains to the processors. In this way the resulting linear systems become smaller compared to the global one. Generally, the two CPU time consuming parts of a finite element program are the calculation of the entries in the system matrices (element loop) and the solving of the resulting linear systems.

In particular, the calculation of the system matrices is done element-wise without any dependency on the computation between different elements. Therefore, the contributions of different elements can be calculated on different processors. When solving the linear system of equations additional care has to be taken for the unknowns residing on the artificial boundaries (fig. 4).

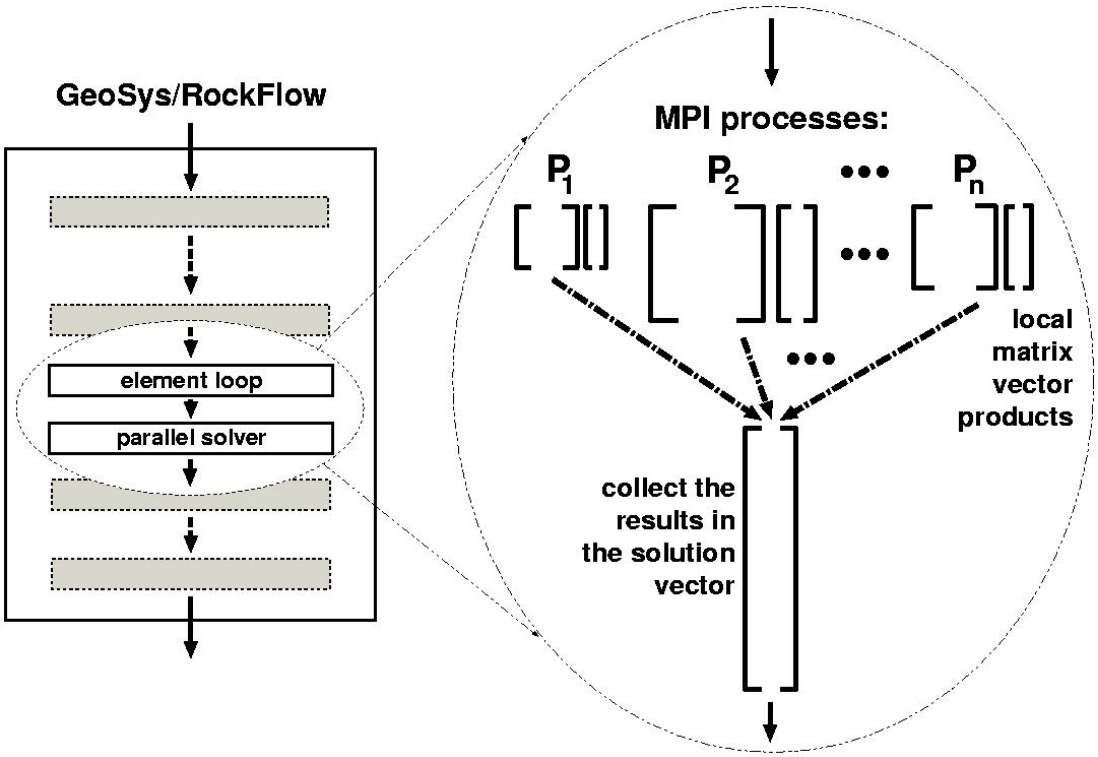

Figure 4: Structure of the future parallel *GeoSys/RockFlow* code

Domain Decomposition and Load Balancing

Load Balancing Tools. When using large parallel computing systems the work load has to be evenly divided among the processing nodes to avoid imbalances among the single computing instances. Known tools to accomplish good load balancing are *Jostle[3], Metis [4]* or *ParMetis*. For our first tests we have chosen Jostle but the remaining two tools, especially ParMetis, a parallel version of Metis, will be tested in future stages of the project as well.

Model Transformation. To solve the problem of load balancing for the real world model it is transformed into a graph on which a partitioning algorithm is applied [5]. The FE mesh is represented in the graph in such a way that the elements of the model are the nodes of the graph and edges only exist between two nodes if the corresponding elements in the mesh are neighbours. For the complete construction of the resulting graph to solve the load balancing problem a neighborhood search is required. For huge models this could be a problem in regard to memory space depending on the underlying hardware.

Implementation. Out of this consideration we have implemented the neighbourhood search using large bitarrays instead of other common data types. Depending on the architecture of the computing environment (if it is a 32 or 64 bit architecture) in one unsigned integer in C at once 32 and 64 different numbers can be stored respectively. This reduces the amount of necessary memory space for the preprocessing of the model data by a factor of 32 or 64. In addition, bit shifting operations are also fast.

Achieved Load Balancing Factors. The resulting graph serves as input for Jostle which separates the original model into almost evenly divided chunks. These partitions of the FE mesh can then be distributed on the computing nodes. In our first tests we have used real models built from either 5528 or 17141 triangular shaped elements. Independent of the number of partitions (from 4

to 12) the load balancing factor that Jostle achieves ranges between 1.012 (4 domains and 17141 elements) as the best result and 1.029 (i.e. 12 domains and 17141 elements) as the worst. Examining the amount of cut edges demonstrates the quality of this domain decomposition: in the best case only 1262, in the worst case also just 3150 edges have been cut from a total of 102541 edges throughout the whole graph. In the future, bigger meshes with different geometric shapes of the elements will be measured too. Figure 5 depicts the way from the real world problem to a parallel computable model.

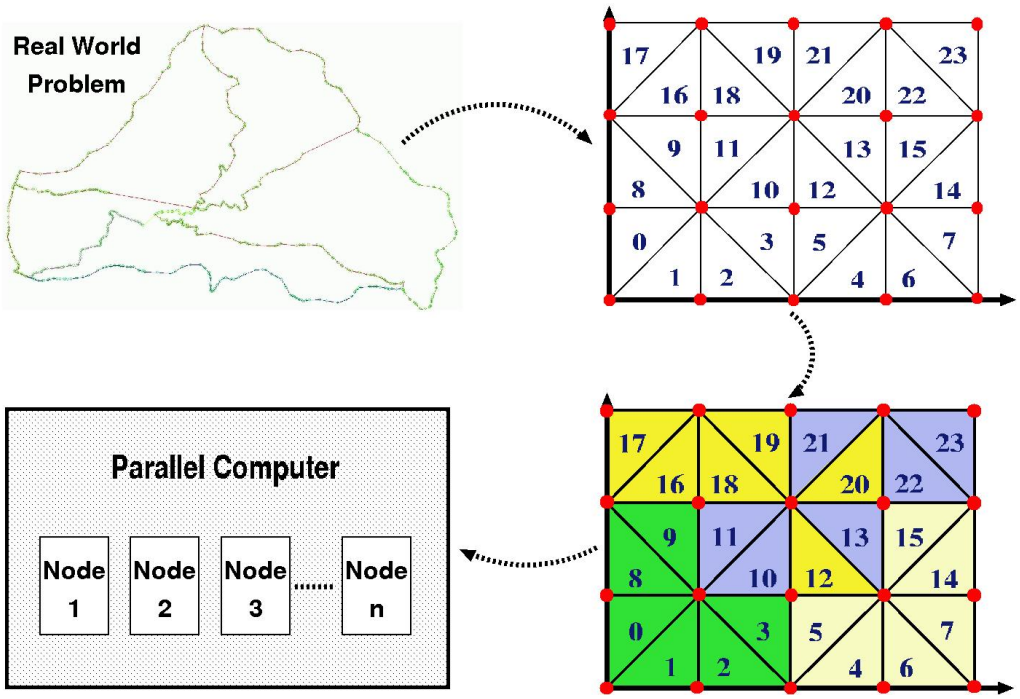

Figure 5: The way from a real world problem to a final computable parallel model

Conclusion

In this work we have presented a geo-process software system, modeling complex coupled processes found in geo-systems. The results of the simulation of a water resources management model for the Jordan Valley show that it is possible to make long-term simulations. But, in order to be more accurate the spatial discretisation has to be increased and the model must include also the solution of other fields, like the temperature field.

This leads to the necessity of using parallel computing systems. In order to achieve this a parallelization strategy based on domain decomposition has been proposed. The first results of the load balancing tool are very satifactory. Further work will be done on parallelizing the solver algorithm alongside with coupling of shallow water and groundwater flow, as well as including kinetic biogeochemical reactions and the improvement of the multi-view and three-dimensional graphics for the GUI.

Acknowledgement

We are very grateful to the Landesstiftung Baden-Wuerttemberg which is funding our project under the number 727.

References

[1] GeoSys/RockFlow Web page. http://www.uni-tuebingen.de/zag/geohydrology/index.html.

[2] Cui Chen, Ali Sawarieh, Thomas Kalbacher, Martin Beinhorn, Wenqing Wang and Olaf Kolditz. Gis based hydrosystem model: Application to the zarqa ma'in - jiza areas in central jordan. Technical report, Center for Applied Geosciences, University of Tübingen, 2004.

[3] C. Walshaw and M. Cross. Parallel Optimisation Algorithms for Multilevel Mesh Partitioning. *Parallel Comput.*, 26(12):1635–1660, 2000. (originally published as Univ. Greenwich Tech. Rep. 99/IM/44).

[4] G. Karypis and V. Kumar. METIS:A Software Package for Partitioning Unstructured Graphs, Partitioning Meshes, and Computing Fill-Reducing Orderings of Sparse Matrices. Technical report, University of Minnesota,Department of Computer Science / Army HPC Research Center, 1998.

[5] C. Walshaw, M. Cross, and M. G. Everett. Parallel Dynamic Graph Partitioning for Adaptive Unstructured Meshes. *J. Parallel Distrib. Comput.*, 47(2):102–108, 1997. (originally published as Univ. Greenwich Tech. Rep. 97/IM/20).