



**A parallelization
scheme to simulate
reactive transport**

W. He et al.

A parallelization scheme to simulate reactive transport in the subsurface environment with OGS#IPhreeqc

W. He^{1,3}, C. Beyer⁴, J. H. Fleckenstein², E. Jang^{1,3}, O. Kolditz^{1,3}, D. Naumov⁵,
and T. Kalbacher¹

¹Department of Environmental Informatics, Helmholtz Centre for Environmental Research – UFZ, Leipzig, Germany

²Department of Hydrogeology, Helmholtz Centre for Environmental Research – UFZ, Leipzig, Germany

³Applied Environmental System Analysis, Technical University Dresden, Dresden, Germany

⁴Institute of Geosciences, Geohydromodeling, Christian-Albrechts-University of Kiel, Kiel, Germany

⁵Faculty of Mechanical and Energy Engineering, Leipzig University of Applied Science, Leipzig, Germany

Received: 15 January 2015 – Accepted: 15 February 2015 – Published: 3 March 2015

Correspondence to: W. He (wenkui.he@ufz.de)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Abstract

This technical paper presents an efficient and performance-oriented method to model reactive mass transport processes in environmental and geotechnical subsurface systems. The open source scientific software packages OpenGeoSys and IPhreeqc have been coupled, to combine their individual strengths and features to simulate thermo-hydro-mechanical-chemical coupled processes in porous and fractured media with simultaneous consideration of aqueous geochemical reactions. Furthermore, a flexible parallelization scheme using MPI (Message Passing Interface) grouping techniques has been implemented, which allows an optimized allocation of computer resources for the node-wise calculation of chemical reactions on the one hand, and the underlying processes such as for groundwater flow or solute transport on the other hand. The coupling interface and parallelization scheme have been tested and verified in terms of precision and performance.

1 Introduction

Reactive transport modeling is an important approach to better understand, quantify and predict hydro-biogeochemical processes and their effects on subsurface environments. It is of growing interest among the fields of geotechnical engineering applications and environmental impact assessments and is used e.g. in contaminated site remediation or water resources management, to predict the environmental fate of organic and inorganic substances and pollutants in soil or groundwater reservoirs (e.g. Ballarini et al., 2014; Hammond et al., 2010, 2011, 2014; Henzler et al., 2014; Lichtner et al., 2012; Molins et al., 2010; Riley et al., 2014; Yabusaki et al., 2011). Geotechnical applications employ reactive transport simulations e.g. to quantify geochemical processes in geological nuclear waste repositories (e.g. Kosakowski and Watanabe, 2013; Shao et al., 2009; Xie et al., 2006) or to evaluate CO₂ geological sequestration (e.g. Beyer et al., 2012; Li et al., 2014; Pau et al., 2010; Xu et al., 2004, 2006).

GMDD

8, 2369–2402, 2015

A parallelization scheme to simulate reactive transport

W. He et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



A parallelization scheme to simulate reactive transport

W. He et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



In the last decades, much effort has been invested to develop practical tools for reactive transport modeling (Steefel et al., 2014), such as PHREEQC (Parkhurst and Appelo, 1999, 2013), OpenGeoSys (OGS) (Kolditz et al., 2012), HYTEC (van der Lee et al., 2003), ORCHESTRA (Meeussen, 2003), TOUGHREACT (Xu and Pruess, 2001; Xu et al., 2006, 2011), eSTOMP (Yabusaki et al., 2011), HYDROGEOCHEM (Yeh and Tripathi, 1990), CrunchFlow (Steefel et al., 2014), MIN3P (Mayer et al., 2002) or PFLO-TRAN (Lichtner et al., 2015). Since each code has its own strengths and limitations, coupling of different codes, i.e. one software applies another and/or vice versa, is an indispensable choice and a straightforward solution to make use of combined capabilities of different codes. Existing approaches, which apply tool coupling methods to simulate reactive transport processes are e.g. HYDRUS and PHREEQC (Jacques and Šimůnek 2005; Šimůnek et al., 2006); COMSOL and PHREEQC (Nardi et al., 2014; Nasir et al., 2014; Wissmeier and Barry, 2011); OGS-GEMs (Kosakowski and Watanabe, 2013; Shao et al., 2009); OGS-BRNS (Centler et al., 2010); OGS-ChemApp (Li et al., 2014); OGS-PHREEQC (Xie et al., 2006; de Lucia et al., 2012); MODFLOW-UFZ and RT3D (Bailey et al., 2013), or MODFLOW-MT3DMS, i.e. PHT3D (Morway et al., 2013).

Due to the complexity of physical, geochemical, and biological processes involved, the development of a reactive transport simulator, which has comprehensive numerical modeling capabilities, is a challenging task. The robustness and computational efficiency of a numerical simulator are of vital importance, because reactive transport modeling is often accompanied with other challenges such as numerical precision and stability (de Dieuleveult and Erhel, 2010; Kosakowski and Watanabe, 2013; Wissmeier and Barry, 2011) or expensive computational time.

Especially for realistic reactive transport simulations at larger scales, i.e. from field scales to catchment or reservoir scale, high complexities of hydrogeological and geochemical systems as well as high spatial–temporal resolution of reactive zones are required to ensure plausible and accurate model results. In these cases, iterative simulations of different scenarios or setups e.g. for model calibration and parameter sensi-

tivity analysis becomes extremely difficult and time-consuming on desktop computers with limited computational resources (Hammond et al., 2014; Kollet et al., 2010; Lichtner et al., 2012; Yabusaki et al., 2011).

Parallelization is an established approach to improve computational performance and with the additional benefit from continuous innovation of modern hardware and software development (Hanappe et al., 2011; Wang et al., 2014). PFLOTRAN, a parallel multiscale and multiphysics code for subsurface multiphase flow and reactive transport (Hammond et al., 2012, 2014; Lichtner et al., 2012), or TOUGH-MP, the parallel version of TOUGH2 (Zhang et al., 2008; Hubschwerlen et al., 2012), apply domain decomposition (DDC) methods for their parallel framework. Yabusaki et al. (2011) implemented a one-sided communication and global shared memory programming paradigm in eSTOMP.

An elaborated code concept and development can help to reduce the time needed for solution procedures and data communication. Consequently in terms of coupled reactive transport modeling, process simulation and interaction should be closely tied to enable shared data structures and reduce data exchange procedures.

In the current work, OGS has been coupled with the new C++ module of PHREEQC, called IPhreeqc. In this operator splitting approach, chemical reactions are calculated locally on each finite element node, whereas processes such as groundwater flow and mass transport are calculated globally. OGS is an open source FEM simulator for multi-dimensional thermo-hydro-mechanical-chemical (THMC) coupled processes in porous and fractured media (Kolditz et al., 2012). In other words, OGS is able to simulate e.g. water and/or gas flow together with heat and mass transport processes in fully and partly saturated media. IPhreeqc on the other hand, inherits all the functionalities of PHREEQC, i.e. it is capable of modelling aqueous, mineral, gas, surface, ion-exchange, solid-solution equilibria and kinetic reactions, but also provides a well-defined set of methods for data transfer and management additionally (Charlton and Parkhurst, 2011). Both codes are open source, i.e. the technical coupling could be realized directly on the code level.

A parallelization scheme to simulate reactive transport

W. He et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



A parallelization scheme to simulate reactive transport

W. He et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



ible groundwater flow, overland flow, density-driven flow, unsaturated flow, two phase as well as multiphase flow. In OGS, transport of components in fluid phases is simulated based on the advection–dispersion equation, while geochemical reactions can be modeled by using internal libraries (e.g. the KinReact module for kinetically controlled biogeochemical-reactions; Ballarini et al., 2014) or external couplings with geochemical solvers (e.g. Xie et al., 2006; Shao et al., 2009; Kosakowski and Watanabe, 2013; Centler et al., 2010; Beyer et al., 2012; Li et al., 2014). The code has already been parallelized using MPI (Wang et al., 2009; Ballarini et al., 2014) and PETSc (Wang et al., 2014). More detailed information relating to OGS development concept, code resources, benchmarking, etc. can be found at <http://www.opengeosys.org/>.

2.2 PHREEQC and IPhreeqc

PHREEQC is one of the most widely used open source geochemical solvers. It provides a variety of geochemical reaction capabilities (Parkhurst and Appelo, 1999, 2013). Beside batch reaction simulations, its current capabilities include inverse and one-dimensional reactive transport modeling. IPhreeqc is a C++ module of PHREEQC which is specially designed for the coupling of PHREEQC with other codes. It provides a well-defined series of methods to interact with a client program (Charlton and Parkhurst, 2011). For example, PHREEQC simulation input data can be prepared as a file or a character string in the client program and executed by PHREEQC with different methods such as *RunFile* or *RunString*. Besides writing selected output results into a file, individual data items at a certain position of the result array can be accessed and returned to the client program by using the *GetSelectedOutputValue* method. More detailed information on IPhreeqc and its data manipulation methods can be found in Charlton and Parkhurst (2011).

2.3 OGS#IPhreeqc interface

In the current study, both source codes, OGS and IPhreeqc are statically linked to allow accesses of all the functionalities of both codes (open source concept). The OGS#IPhreeqc interface is well encapsulated into a general framework for reactive transport modeling in OGS, which has already been described in detail by Beyer et al. (2012). Unlike the previously existing coupling scheme between OGS and PHREEQC presented by Xie et al. (2006), in which the PHREEQC is called externally through a system call to a PHREEQC binary executable, in the new coupling presented here, a call to PHREEQC can be realized directly by accessing functions provided by the IPhreeqc module. The source code of PHREEQC however is not changed, which allows the merging of new releases from both codes rather conveniently. This development concept allows the user to benefit continuously from the code development from both sides. The sequential non-iterative (SNIA) approach for operator splitting is applied in the coupling procedure. Figure 1 illustrates the general procedure for reactive transport modeling with OGS#IPhreeqc, which is described in the following.

In the first development step, a file-based approach for data exchange between OGS and IPhreeqc is applied. Before entering the time stepping loop, initial values of the system state such as component concentrations and temperatures on each finite element node will be passed to IPhreeqc to initialize the geochemical system. During each time step, after OGS has calculated the flow field by simulating different flow processes mass transport of each mobile chemical component will be calculated. Then on each node, concentration values of each component as well as other state variables such as pressure and temperature will be forwarded to the coupling interface, in which an input file for IPhreeqc will be prepared. The call to IPhreeqc will be realized by using the IPhreeqc functions: *CreateIPhreeqc* to create a new instance of IPhreeqc, *LoadDatabase* to load the thermodynamic database for the geochemical system, and *RunFile* to run the specific PHREEQC input files. After execution of IPhreeqc, an output file will be generated by IPhreeqc, which will be read by the interface during the reac-

GMDD

8, 2369–2402, 2015

A parallelization scheme to simulate reactive transport

W. He et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



tion post-processing. Based on the updated chemical species concentrations, several feedback functions can be applied to update the porosity, permeability, saturation as well as density for flow, heat and mass transport processes. For example, in the case of mineral dissolution or precipitation, the porosity and permeability changes can be evaluated.

2.4 Verification of the coupling interface

The coupling between OGS and IPhreeqc was tested and verified by using several benchmarks for reactive transport problem types such as ion exchange (example 11 of Parkhurst and Appelo, 1999), carbonate mineral precipitation and dissolution (Engesgaard and Kipp, 1992; Beyer et al., 2012), and isotope fractionation (van Breukelen et al., 2005). The latter two benchmarks will be shortly introduced here.

The first presented test example is the Engesgaard benchmark. It describes the phenomenon occurs when a 0.5 m long 1-D calcite column is flushed with a solution containing magnesium: calcite dissolves continuously as the solution moves towards the downstream direction, whereas dolomite precipitates temporarily at the calcite dissolution front. Calcite dissolution/precipitation are simulated as equilibrium reactions, whereas that of the dolomite is modelled as kinetic reactions using the rate parameters from Palandri and Kharaka (2004). The material properties of the column as well as the initial and boundary conditions are given in Tables 1 and 2, respectively. The comparison of the simulation results between OGS#IPhreeqc and OGS-Chemapp (from Beyer et al., 2012) shows a very good agreement as illustrated in Fig. 2.

The second benchmark is based on the 1-D multistep isotope fractionation model from van Breukelen et al. (2005), which simulates the sequential reductive dechlorination of tetrachloroethene (PCE) to ethane (ETH) in a 876 m long aquifer over a period of 20 years. The model domain, aquifer properties as well as initial and boundary conditions are illustrated in Fig. 3.

A parallelization scheme to simulate reactive transport

W. He et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



The intermediate products during the degradation include tri- and dichloroethylene (TCE, DCE), vinyl chloride (VC). The whole sequential reductive dechlorination chain is illustrated as follows: $PCE \rightarrow TCE \rightarrow DCE \rightarrow VC \rightarrow ETH$.

The ^{12}C and ^{13}C isotopes of each chlorinated hydrocarbons (CHCs) are modeled as separate species. Totally there are 11 chemical species including chloride as tracer, which is produced in each dechlorination reaction. During degradation the kinetic isotope fractionation of each compound is assumed to be constant. More detailed information regarding to the kinetic rate expressions and relevant parameters can be found in van Breukelen et al. (2005).

The simulated concentration profile of the light CHC isotopes and relevant $\delta^{13}C$ [‰] isotope signatures along the model domain are compared with those simulated using a standalone version of PHREEQC (Fig. 4), showing a good agreement for both concentration profiles of the light CHC isotopes and corresponding isotope signatures.

3 Parallelization of OGS#IPhreeqc

In this section we describe the parallelization method for the numerical simulation of reactive transport processes with OGS#IPhreeqc. For the parallelization of groundwater flow and mass transport, the OGS internal DDC scheme (see Sect. 2.1) is employed. For the parallelization of geochemical reactions a loop parallelization is applied. All cores take part in solving the geochemical reaction system, while only certain cores are used to solve the DDC related processes.

3.1 Parallelization scheme

Figures 5 and 6 illustrate the general idea of the parallelization scheme. The two different MPI groups, i.e. MPI_Group1 and MPI_Group2 and related intra-communicators are created by using MPI functions MPI_Group_incl and MPI_Comm_create. The compute cores which belong to MPI_Group1 will run most part of the OGS code including

A parallelization scheme to simulate reactive transport

W. He et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



all DDC related processes (groundwater flow, mass and heat transport) and geochemical reactions, whereas those of MPI_Group2 will only run a small part of code related to geochemical simulation.

A “for” loop for MPI_Group2 is created directly in the main function of the OGS code. In each time step, after the calculation of global flow and mass transport process, PHREEQC input files for all compute cores will be created by compute cores of MPI_Group1. Then MPI_Group1 will send a signal to MPI_Group2, which will invoke the calls to IPhreeqc for compute cores in MPI_Group2. After PHREEQC calculations are complete in both MPI groups, flow and mass transport processes will start again with the next time step in MPI_Group1, while compute cores of MPI_Group2 will wait for the signal from MPI_Group1 (using blocking receive MPI_Receive) to restart the calls to IPhreeqc. After compute cores of MPI_Group1 have run through the complete time stepping loop reaching the end of the simulation, another signal will be sent to MPI_Group2, which will force its compute cores to jump out of the chemical reaction loops. Then MPI_Finalize will be executed to terminate the MPI environment. As a special case, when the number of subdomains equals that of the compute cores, MPI_Group2 will not be created. In this case, no communication between the two MPI groups is required. This corresponds to the parallelization scheme for reactive transport simulations applied in Ballarini et al. (2014).

As mentioned above, file-based data transfer is applied to exchange concentration values between mass transport and geochemical reaction simulations. In each time step, after the simulation of mass transport, concentration values of all components in all finite element nodes will be stored in a global concentration vector. For each compute core a node list vector will be generated through which finite element nodes are allocated to the respective compute core, and their concentration values can be accessed from the global concentration data structure by using this vector. Since the generation of the node list vector is completely independent from the domain decomposition, flexible groupings of finite element nodes can be realized to ensure an optimum load balance of compute cores for the calculation of geochemical reactions. During the

A parallelization scheme to simulate reactive transport

W. He et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



execution of geochemical reactions, each compute core will execute IPhreeqc by using a specific input file. A relevant PHREEQC results file will then be generated. After all compute cores finish their calls to IPhreeqc, compute cores of MPI_Group1 will read all the result files and store the concentration values of all components in respective local buffers. The values of all local buffers will then be transferred to a global concentration vector by applying the MPI_Allgather method, before the updated concentrations of different components are sent back to mass transport process again.

3.2 Computational platforms

The correctness and efficiency of the proposed scheme were tested on two different computational platforms. The first platform is a multi-core Linux machine called “ENVINF”. It contains 40 “Intel(R) Xeon(R) CPU E5-2680 v2 @ 2.80 GHz” CPU cores and has a shared memory of approximately 500 GB RAM among these 40 cores. A maximum of 20 cores can be used by a single user at a time. The second platform is a Linux based (CentOS 6 as the operating system) cluster, in the following called “EVE”. It consists of 1008 (Intel XEON X5650 @ 2.6 GHz) CPU cores and 5.5 TB of RAM. Computer nodes are connected with 40 Gbits⁻¹ QDR Infiniband network interconnect. The peak performance is 10 TFlops⁻¹.

In order to make the results comparable by using both platforms, for all tests in the EVE cluster, job requests were made to guarantee the use of compute nodes with 20 free slots when submitting to the job queue. Of course jobs can also be submitted without this constrain, however, since in this case the MPI jobs may be distributed to more compute nodes than necessary in order to allow an earlier execution, more inter-compute node communications may have to be made over the network, which would worsen the performance of the parallelization scheme.

GMDD

8, 2369–2402, 2015

A parallelization scheme to simulate reactive transport

W. He et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



different DDCs anymore. Curve AB in Fig. 8a represents relative speedups for combinations in which the number of compute cores equals the number of DDCs. In Fig. 8b curve AB is once again illustrated (“total”) together with the speedup of chemical reactions and speedup of “other” processes which are mainly global processes like ground-water flow and mass transport. We can observe that the speedup of “other” processes reaches its maximum when the number of compute cores exceeds 12. As shown by Wang et al. (2009), adding of sub-domains will increase communication between sub-domain border nodes. As a consequence, the parallel efficiency for calculation of DDC related processes will reduce when number of border nodes becomes comparable with the total number of finite element nodes. The speedup of reaction however is generally much better and increases continuously as more compute cores are provided. In the operator splitting approach chemical reactions are solved locally on each finite element node, hence no direct communication among different nodes is necessary.

Figure 8c and d show the breakdown of time for different processes with a DDC = 4 and a DDC = 12. It is clearly shown that chemical reaction is the most time-consuming part of the simulation in both cases. With a DDC = 4 reactions take up to 90 % of the total time when only 4 compute cores are applied, and drops to around 70 % if 20 compute cores are applied; whereas for a DDC = 12 it becomes 83 % of the total time for 12 compute cores, and goes down to around 77 % for 20 compute cores. In both cases time of “other” processes stays almost unchanged for different number of compute cores because the number of DDCs is fixed. The time for interface mainly includes preparing input files for IPhreeqc, communication among different compute cores, reading output files from IPhreeqc. As shown in Fig. 8c and d, this part of time stays also nearly unchanged and relatively low compared to other processes. Generally, the way of coupling and parallelization is shown to be efficient already for small sized reactive transport problems in a shared memory system such as ENVINF.

GMDD

8, 2369–2402, 2015

A parallelization scheme to simulate reactive transport

W. He et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)



[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



4.2 Isotope fractionation 3-D

The second test case is a 3-D extension (876 m × 100 m × 10 m, see Fig. 7b) of the 2-D test example which consists of 134 431 nodes and 120 000 hexahedral finite elements (120 × 100 × 10). The simulation time with 2 compute cores with 2 DDCs on ENVINF is 37.5 h.

Similar to the 2-D test example (Sect. 4.1), for the 3-D test case the relative speedup on the EVE cluster is illustrated as a function of number of DDCs and total compute cores in Fig. 9a; Fig. 9b shows a breakdown of curve AB into speedups of global processes and chemical reactions. If we use the same number of compute cores and DDCs, a nearly linear speedup with the increase of the compute cores can be observed. By using 80 compute cores simulation time can be reduced to around 45 min. As problem size increases, the speedup effects of both DDC related processes as well as chemical reactions become stronger. Similar to the results of the 2-D example, in the 3-D example geochemical reaction shows a much better speedup than global processes.

However, if we fix the number of DDCs at a specific value and increase the total compute cores further, there is not much speedup observed for almost all DDCs (see Fig. 9a). This behavior is quite different from what we have observed in the 2-D example.

The reason behind lies mainly in the fact, that the ratio between the time consumption for reactions and mass transport (flow) are different in these two examples. In the 2-D example, the time consumption for calculation of flow and mass transport is rather low comparing with that of reactions. In the 3-D example, the time consumption for flow and mass transport is on the same magnitude as that of reactions (see Fig. 10a and b). As a consequence, the saving of time in the calculation of reactions alone, which is obtained by increasing compute cores, cannot bring a significant speedup for the entire simulation. Hence, to achieve a better speedup for a large problem, it is important to reduce the time consumption for flow and mass transport as well by using more DDCs.

GMDD

8, 2369–2402, 2015

A parallelization scheme to simulate reactive transport

W. He et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



In the current scheme the writing and the reading of files on the EVE cluster is realized by using the general parallel file system (GPFS). As we can observe in the Fig. 10c, for a large problem file writing and reading through the GPFS is a time-consuming procedure. Figure 10d shows that other time consumption in the coupling interface is relatively small, which mainly consists of blocking communication (e.g. MPI_Barrier) between different compute cores. However, this part of time can also increase when working load for calculation of reactions, file writing and reading are unbalanced among different compute cores.

5 Conclusions and outlook

This technical paper introduced the coupling interface OGS#IPhreeqc and a parallelization scheme developed for the interface. Furthermore, the parallel performance of the scheme was analyzed.

The OGS#IPhreeqc interface presented in the current study is a sustainable way of coupling, which can continuously benefit from the code development and updating from two open source communities. The parallelization scheme is adjustable to different hardware architectures, and suitable for different types of high performance computing (HPC) platforms such as shared-memory machines or clusters.

The parallelization scheme provides more flexibility to arrange compute resources for different computational tasks by using the MPI grouping concept. The appropriate setting of DDCs and total compute cores is problem dependent. If calculation of geochemical reaction dominates the total simulation time, e.g. for small sized problems with simple hydrogeological but complex geochemical system, then using more compute cores brings more significant speedup than simply increasing the number of DDCs; if the time consumption of flow and mass transport is in the same magnitude as geochemical reactions, to increase the number of compute cores and DDCs simultaneously is most efficient.

A parallelization scheme to simulate reactive transport

W. He et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



A parallelization scheme to simulate reactive transport

W. He et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



The current parallelization scheme will be especially useful for problems, in which a further increase of the number of DDCs above the optimum will lead to a strong degradation of parallel performance for flow or mass transport (as a consequence of e.g. increasing inter-compute-node communication or number of linear iterations). In this case, better speedup may still be obtained by fixing the number of DDCs at the optimum while increasing the number of compute cores to accelerate the calculation of chemical reactions.

Even though the current parallelization scheme has shown good parallel performance in shared- and distributed-memory systems, there is still space for improvement to reduce the time consumption for communication and data transfer between OGS and IPhreeqc.

Blocking communication techniques, like MPI_Barrier were applied to ensure the correct sequence of process coupling. An unbalanced work load distribution for chemical reactions, like in heterogeneous problems with sharp transient reactive fronts or reaction hot spots, could affect the parallel performance as well. Hence, more intelligent ways to ensure efficient load balance still remain as an important task.

File writing and reading through the GPFS of a cluster system is time-consuming, especially when increasing problem size. As the next step, a character string-based data exchange strategy will be implemented, in order to minimize the time consumption for data exchange.

Recently, the SeS Bench (Subsurface Environmental Simulation Benchmarking) benchmarking initiative has started a project to test the parallel performance of different reactive transport modeling tools. In the near future, more complex benchmarks and real-world applications will be tested in the framework of this project to improve the parallel performance of the current scheme for reactive transport modeling at larger scales.

Acknowledgements. This work is funded by Helmholtz Centre for Environmental Research (UFZ), Leipzig, Germany. The authors thank the relevant Integrated Project T31 “Catchment Dynamics” of POF3 and its coordinator Maren Göhler. The authors thank Lars Bilke for

his technical supports for the coupling procedure and Ben Langenberg for helps in running simulations on the EVE cluster.

The service charges for this open-access publication
5 have been covered by a Research Centre of the
Helmholtz Association.

References

Bailey, R. T., Morway, E. D., Niswonger, R. G., and Gates, T. K.: Modeling variably saturated multispecies reactive groundwater solute transport with MODFLOW-UZF and RT3D, *Ground Water*, 51, 752–761, 2013.

Ballarini, E., Beyer, C., Bauer, R. D., Griebler, C., and Bauer, S.: Model based evaluation of a contaminant plume development under aerobic and anaerobic conditions in 2-D bench-scale tank experiments, *Biodegradation*, 25, 351–371, 2014.

Beyer, C., Li, D., De Lucia, M., Kühn, M., and Bauer, S.: Modelling CO₂-induced fluid-rock interactions in the Altensalzwedel gas reservoir. Part II: coupled reactive transport simulation, *Environ. Earth Sci.*, 67, 573–588, 2012.

Centler, F., Shao, H. B., De Biase, C., Park, C. H., Regnier, P., Kolditz, O., and Thullner, M.: GeoSysBRNS – A flexible multidimensional reactive transport model for simulating biogeochemical subsurface processes, *Comput. Geosci.*, 36, 397–405, 2010.

Charlton, S. R. and Parkhurst, D. L.: Modules based on the geochemical model PHREEQC for use in scripting and programming languages, *Comput. Geosci.*, 37, 1653–1663, 2011.

de Dieuleveult, C. and Erhel, J.: A global approach to reactive transport: application to the MoMas benchmark, *Comput. Geosci.*, 14, 451–464, 2010.

de Lucia, M., Bauer, S., Beyer, C., Kühn, M., Nowak, T., Pudlo, D., Reitenbach, V., and Stadler, S.: Modelling CO₂-induced fluid-rock interactions in the Altensalzwedel gas reservoir. Part I: from experimental data to a reference geochemical model, *Environ. Earth Sci.*, 67, 563–572, 2012.

Engesgaard, P. and Kipp, K. L.: A geochemical transport model for redox-controlled movement of mineral fronts in groundwater-flow systems – a case of nitrate removal by oxidation of pyrite, *Water Resour. Res.*, 28, 2829–2843, 1992.

GMDD

8, 2369–2402, 2015

A parallelization scheme to simulate reactive transport

W. He et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



A parallelization scheme to simulate reactive transport

W. He et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Hammond, G. E. and Lichtner, P. C.: Field-scale model for the natural attenuation of uranium at the Hanford 300 Area using high-performance computing, *Water Resour. Res.*, 46, W09527, doi:10.1029/2009WR008819, 2010.

Hammond, G. E., Lichtner, P. C., and Rockhold, M. L.: Stochastic simulation of uranium migration at the Hanford 300 Area, *J. Contam. Hydrol.*, 120–121, 115–128, doi:10.1016/j.jconhyd.2010.04.005, 2011.

Hammond, G. E., Lichtner, P. C., and Mills, R. T.: Evaluating the performance of parallel subsurface simulators: an illustrative example with PFLOTTRAN, *Water Resour. Res.*, 50, 208–228, 2014.

Hanappe, P., Beurivé, A., Laguzet, F., Steels, L., Bellouin, N., Boucher, O., Yamazaki, Y. H., Aina, T., and Allen, M.: FAMOUS, faster: using parallel computing techniques to accelerate the FAMOUS/HadCM3 climate model with a focus on the radiative transfer algorithm, *Geosci. Model Dev.*, 4, 835–844, doi:10.5194/gmd-4-835-2011, 2011.

Henzler, A. F., Greskowiak, J., and Massmann, G.: Modeling the fate of organic micropollutants during river bank filtration (Berlin, Germany), *J. Contam. Hydrol.*, 156, 78–92, 2014.

Hubschwerlen, N., Zhang, K. N., Mayer, G., Roger, J., and Vialay, B.: Using Tough2-MP on a cluster-optimization methodology and study of scalability, *Comput. Geosci.*, 45, 26–35, 2012.

Jacques, D. and Šimůnek, J.: User Manual of the Multicomponent Variably-saturated Flow and Transport Model HP1, Description, Verification and Examples, Version 1.0, SCK-CEN-BLG-998, Waste and Disposal, SCK-CEN, Mol, Belgium, 79 pp., 2005.

Kolditz, O., Bauer, S., Bilke, L., Bottcher, N., Delfs, J. O., Fischer, T., Gorke, U. J., Kalbacher, T., Kosakowski, G., McDermott, C. I., Park, C. H., Radu, F., Rink, K., Shao, H., Shao, H. B., Sun, F., Sun, Y. Y., Singh, A. K., Taron, J., Walther, M., Wang, W., Watanabe, N., Wu, Y., Xie, M., Xu, W., and Zehner, B.: OpenGeoSys: an open-source initiative for numerical simulation of thermo-hydro-mechanical/chemical (THM/C) processes in porous media, *Environ. Earth Sci.*, 67, 589–599, 2012.

Kollet, S. J., Maxwell, R. M., Woodward, C. S., Smith, S., Vanderborght, J., Vereecken, H., and Simmer, C.: Proof of concept of regional scale hydrologic simulations at hydrologic resolution utilizing massively parallel computer resources, *Water Resour. Res.*, 46, W04201, doi:10.1029/2009WR008730, 2010.

**A parallelization
scheme to simulate
reactive transport**

W. He et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Kosakowski, G. and Watanabe, N.: OpenGeoSys-Gem: a numerical tool for calculating geochemical and porosity changes in saturated and partially saturated media, *Phys. Chem. Earth*, 70–71, 138–149, 2014.
- Li, D., Bauer, S., Benisch, K., Graupner, B., and Beyer, C.: OpenGeoSys-ChemApp: a coupled simulator for reactive transport in multiphase systems and application to CO₂ storage formation in Northern Germany, *Acta. Geotech.*, 9, 67–79, 2014.
- Lichtner, P. C. and Hammond, G. E.: Using high performance computing to understand roles of labile and nonlabile uranium(VI) on Hanford 300 Area Plume Longevity, *Vadose Zone J.*, 11, doi:10.2136/vzj2011.0097, 2012.
- Lichtner, P. C., Hammond, G. E., Lu, C., Karra, S., Bisht, G., Andre, B., Mills, R. T., and Kumar, J.: PFLOTRAN User Manual: A Massively Parallel Reactive Flow and Transport Model for Describing Surface and Subsurface Processes, available at: http://www.pflotran.org/docs/user_manual.pdf, last access: 2 March 2015.
- Mayer, K. U., Frind, E. O., and Blowes, D. W.: Multicomponent reactive transport modeling in variably saturated porous media using a generalized formulation for kinetically controlled reactions, *Water Resour. Res.*, 38, 1174, doi:10.1029/2001WR000862, 2002.
- Meeussen, J. C. L.: ORCHESTRA: an object-oriented framework for implementing chemical equilibrium models, *Environ. Sci. Technol.*, 37, 1175–1182, 2003.
- Molins, S., Mayer, K. U., Amos, R. T., and Bekins, B. A.: Vadose zone attenuation of organic compounds at a crude oil spill site – interactions between biogeochemical reactions and multicomponent gas transport, *J. Contam. Hydrol.*, 112, 15–29, 2010.
- Morway, E. D., Niswonger, R. G., Langevin, C. D., Bailey, R. T., and Healy, R. W.: Modeling variably saturated subsurface solute transport with MODFLOW-UZF and MT3DMS, *Ground Water*, 51, 237–251, 2013.
- Nardi, A., Idiart, A., Trincherro, P., de Vries, L. M., and Molinero, J.: Interface COMSOL-PHREEQC (iCP), an efficient numerical framework for the solution of coupled multiphysics and geochemistry, *Comput. Geosci.*, 69, 10–21, 2014.
- Nasir, O., Fall, M., and Evgin, E.: A simulator for modeling of porosity and permeability changes in near field sedimentary host rocks for nuclear waste under climate change influences, *Tunn. Undergr. Sp. Tech.*, 42, 122–135, 2014.
- Palandri, J. L. and Kharaka, Y. K.: A compilation of rate parameters of water-mineral interaction kinetics for application to geochemical modelling, US Geological Survey Water-Resources Investigations Report 04–1068, 2004.

**A parallelization
scheme to simulate
reactive transport**W. He et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

Parkhurst, D. L. and Appelo, C. A. J.: User's guide to PHREEQC (Version 2) – a computer program for speciation, batch-reaction, one-dimensional transport and inverse geochemical calculations, US Geological Survey Water-Resources Investigations Report, 99–4259, 312 pp., 1999.

5 Parkhurst, D. L. and Appelo, C. A. J.: Description of input and examples for PHREEQC version 3 – a computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations, in: US Geological Survey Techniques and Methods, book 6, chap. A43, 497 pp., 2013.

10 Pau, G. S. H., Bell, J. B., Pruess, K., Almgren, A. S., Lijewski, M. J., and Zhang, K. N.: High-resolution simulation and characterization of density-driven flow in CO₂ storage in saline aquifers, *Adv. Water Resour.*, 33, 443–455, 2010.

Riley, W. J., Maggi, F., Kleber, M., Torn, M. S., Tang, J. Y., Dwivedi, D., and Guerry, N.: Long residence times of rapidly decomposable soil organic matter: application of a multi-phase, multi-component, and vertically resolved model (BAMS1) to soil carbon dynamics, *Geosci. Model Dev.*, 7, 1335–1355, doi:10.5194/gmd-7-1335-2014, 2014.

15 Shao, H. B., Dmytrieva, S. V., Kolditz, O., Kulik, D. A., Pfingsten, W., and Kosakowski, G.: Modeling reactive transport in non-ideal aqueous-solid solution system, *Appl. Geochem.*, 24, 1287–1300, 2009.

20 Šimůnek, J., Jacques, D., van Genuchten, M. T., and Mallants, D.: Multicomponent geochemical transport modeling using HYDRUS-1D and HP1, *J. Am. Water Resour. As.*, 42, 1537–1547, 2006.

25 Steefel, C. I., Appelo, C. A. J., Arora, B., Jacques, D., Kalbacher, T., Kolditz, O., Lagneau, V., Lichtner, P. C., Mayer, K. U., Meeussen, J. C. L., Molins, S., Moulton, D., Shao, H., Šimůnek, J., Spycher, N., Yabusaki, S. B., and Yeh, G. T.: Reactive transport codes for subsurface environmental simulation, *Comput. Geosci.*, 1–34, doi:10.1007/s10596-014-9443-x, 2014.

van Breukelen, B. M., Hunkeler, D., and Volkering, F.: Quantification of sequential chlorinated ethene degradation by use of a reactive transport model incorporating isotope fractionation, *Environ. Sci. Technol.*, 39, 4189–4197, 2005.

30 van der Lee, J., De Windt, L., Lagneau, V., and Goblet, P.: Module-oriented modeling of reactive transport with HYTEC, *Comput. Geosci.*, 29, 265–275, 2003.

A parallelization scheme to simulate reactive transport

W. He et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Wang, W., Fischer, T., Zehner, B., Böttcher, N., Görke, U.-J., and Kolditz, O.: A parallel finite element method for two-phase flow processes in porous media: OpenGeoSys with PETSc, *Environ. Earth Sci.*, 73, 2269–2285, doi:10.1007/s12665-014-3576-z, 2014.

Wang, W. Q., Kosakowski, G., and Kolditz, O.: A parallel finite element scheme for thermo-hydro-mechanical (THM) coupled problems in porous media, *Comput. Geosci.*, 35, 1631–1641, 2009.

Wang, W. Q., Rutqvist, J., Gorke, U. J., Birkholzer, J. T., and Kolditz, O.: Non-isothermal flow in low permeable porous media: a comparison of Richards' and two-phase flow approaches, *Environ. Earth Sci.*, 62, 1197–1207, 2011.

Wissmeier, L. and Barry, D. A.: Simulation tool for variably saturated flow with comprehensive geochemical reactions in two- and three-dimensional domains, *Environ. Modell. Softw.*, 26, 210–218, 2011.

Xie, M. L., Bauer, S., Kolditz, O., Nowak, T., and Shao, H.: Numerical simulation of reactive processes in an experiment with partially saturated bentonite, *J. Contam. Hydrol.*, 83, 122–147, 2006.

Xu, T. F. and Pruess, K.: Modeling multiphase non-isothermal fluid flow and reactive geochemical transport in variably saturated fractured rocks: 1. Methodology, *Am. J. Sci.*, 301, 16–33, 2001.

Xu, T. F., Apps, J. A., and Pruess, K.: Numerical simulation of CO₂ disposal by mineral trapping in deep aquifers, *Appl. Geochem.*, 19, 917–936, 2004.

Xu, T. F., Sonnenthal, E., Spycher, N., and Pruess, K.: TOUGHREACT – a simulation program for non-isothermal multiphase reactive geochemical transport in variably saturated geologic media: applications to geothermal injectivity and CO₂ geological sequestration, *Comput. Geosci.*, 32, 145–165, 2006.

Xu, T. F., Spycher, N., Sonnenthal, E., Zhang, G. X., Zheng, L. E., and Pruess, K.: TOUGHREACT version 2.0: a simulator for subsurface reactive transport under non-isothermal multiphase flow conditions, *Comput. Geosci.*, 37, 763–774, 2011.

Yabusaki, S. B., Fang, Y. L., Williams, K. H., Murray, C. J., Ward, A. L., Dayvault, R. D., Waichler, S. R., Newcomer, D. R., Spane, F. A., and Long, P. E.: Variably saturated flow and multicomponent biogeochemical reactive transport modeling of a uranium bioremediation field experiment, *J. Contam. Hydrol.*, 126, 271–290, 2011.

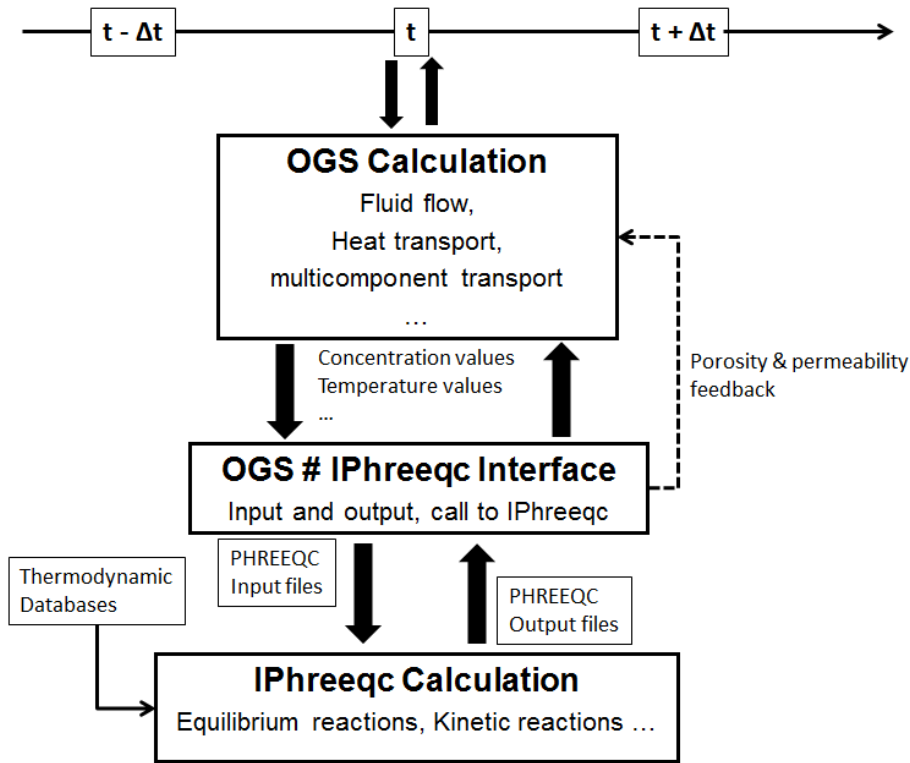


Figure 1. General concept of the coupling interface between OGS and IPhreeqc.

A parallelization scheme to simulate reactive transport

W. He et al.

Title Page	
Abstract	Introduction
Conclusions	References
Tables	Figures
◀	▶
◀	▶
Back	Close
Full Screen / Esc	
Printer-friendly Version	
Interactive Discussion	



A parallelization scheme to simulate reactive transport

W. He et al.

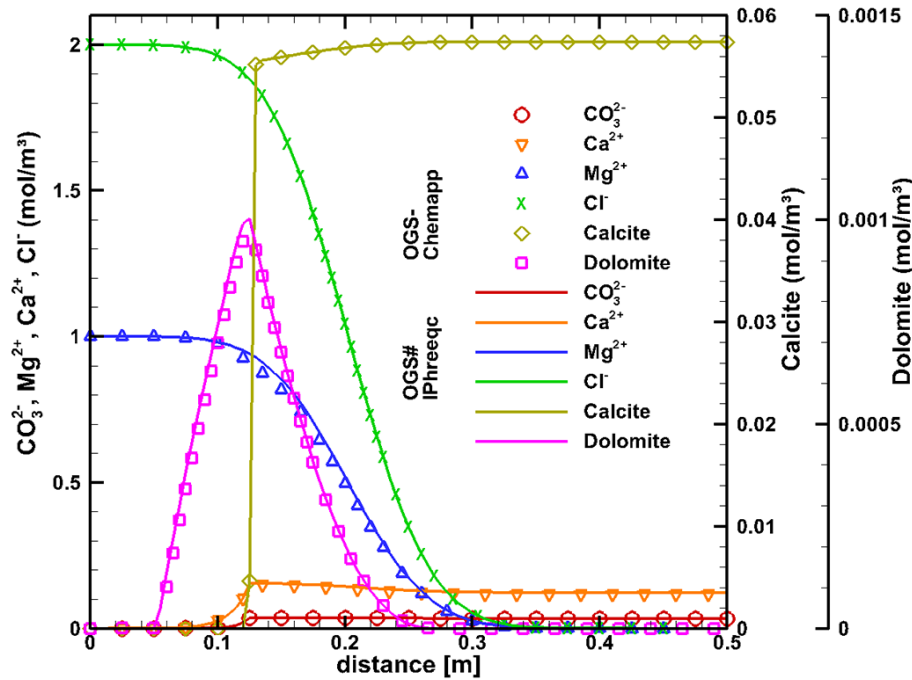


Figure 2. Comparison of calcite and dolomite precipitation/dissolution simulation with OGS-Chemapp and OGS#IPhreeqc.

[Title Page](#)
[Abstract](#) [Introduction](#)
[Conclusions](#) [References](#)
[Tables](#) [Figures](#)
 ◀ ▶
 ◀ ▶
[Back](#) [Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)



A parallelization scheme to simulate reactive transport

W. He et al.

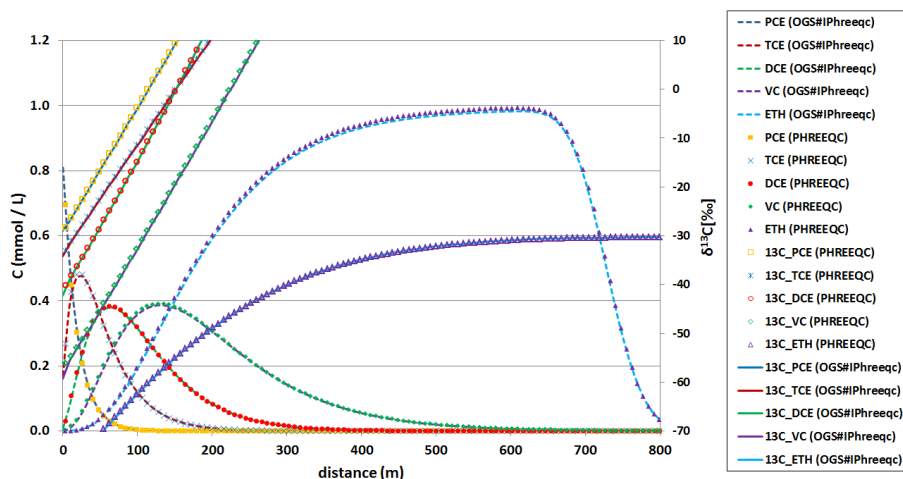


Figure 4. Concentration profiles of the light CHC isotopologues and $\delta^{13}\text{C}$ [‰] isotope signatures along the horizontal axis of the model domain simulated by OGS#IPhreeqc (dashed lines or full lines) and PHREEQC (symbols) at the end of the simulations after 20 years.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



A parallelization scheme to simulate reactive transport

W. He et al.

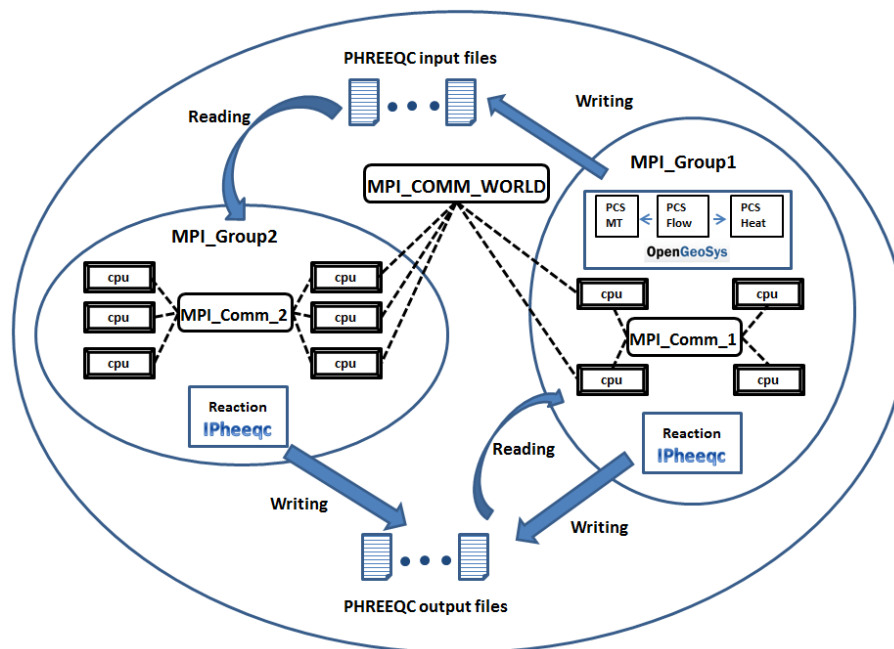


Figure 5. Parallelization scheme for OGS#IPhreeqc. Two distinct MPI groups and relevant inter- and intra-communicators are created. MPI_Group1 take part in the simulation of both DDC related processes and chemical reactions, while MPI_Group2 only participates in the simulation of chemical reactions. PCS MT, PCS Flow and PCS Heat are process of mass transport, flow and heat transport, respectively.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



A parallelization scheme to simulate reactive transport

W. He et al.

```

for (ranks of MPI_Group1)//MPI_Group1 will run following code
{
  read OGS input data;
  while (the current time is smaller than the end time) //time stepping loop
  {
    compute flow processes;
    compute heat and mass transport process;
    prepare the input files for PHREEQC;
    send signals to MPI_Group2;
    //inform MPI_Group2 that the input files for IPhreeqc are prepared;
    calculate chemical reactions with IPhreeqc;
    waiting for signals from MPI_Group2;
    read output files from IPhreeqc;
  }
  output results; //if needed
  send MPI_Group2 a kill signal;
  //inform MPI_Group2 the time stepping loop is over
  terminate MPI environment;
}

for (ranks of MPI_Group2)//MPI_Group2 will run following code
{
  for () //reaction loop
  {
    waiting for signal from MPI_Group1;
    if the signal is a kill signal
      jump out of the reaction loop;
    else
      calculate chemical reactions with IPhreeqc;
      send signals to MPI_Group1;
      //inform MPI_Group1 that calculation of reaction is done
  }
  terminate MPI environment;
}

```

Figure 6. Pseudo code for schematic presentation of the parallelization scheme.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



A parallelization scheme to simulate reactive transport

W. He et al.

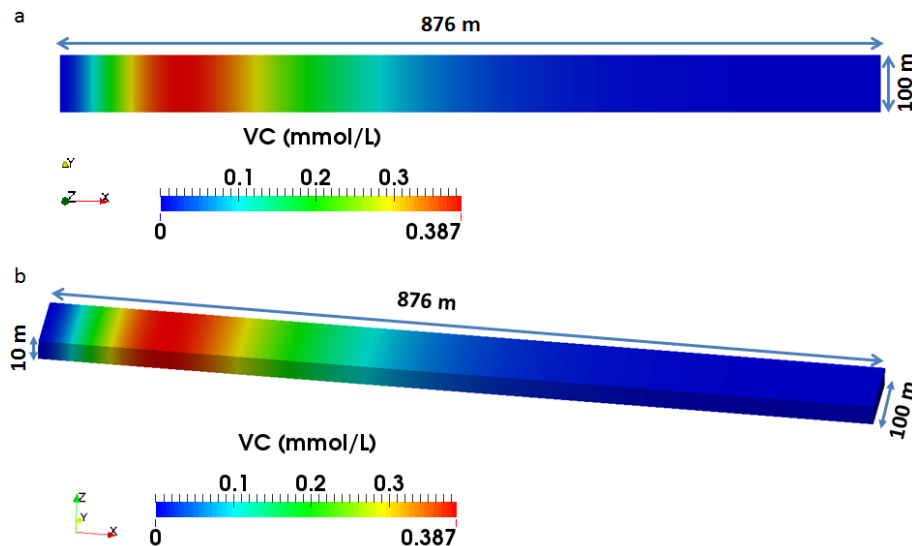


Figure 7. Concentration profile of light isotope VC of the 2-D model (a) and the 3-D model (b) at the end of the simulation. For (b) a vertical (z direction) exaggeration of 2 times is applied.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

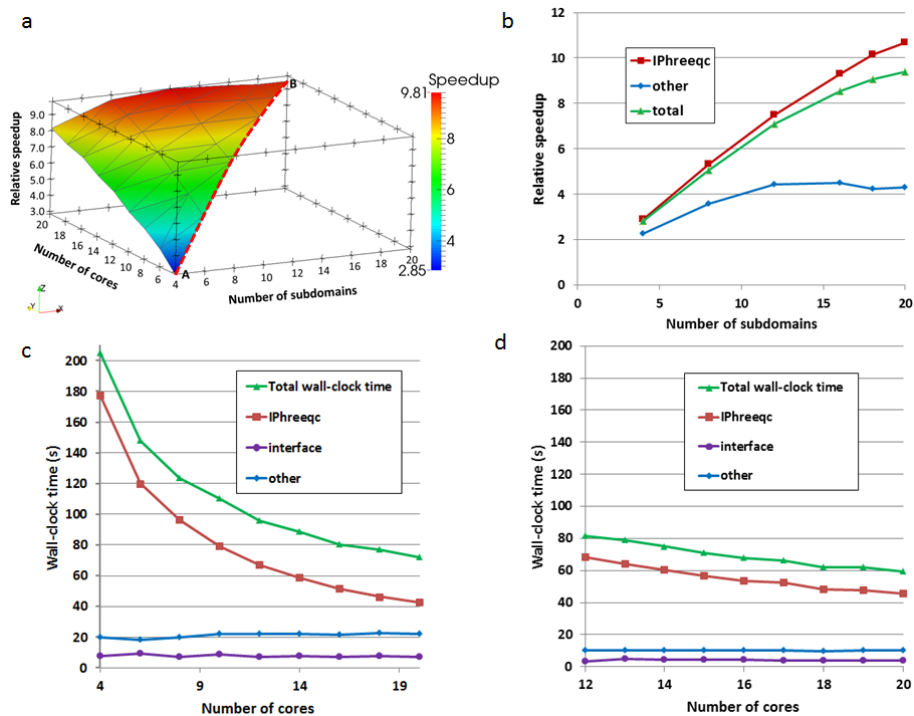


Figure 8. Performance of the proposed parallelization scheme in running isotope fractionation 2-D example on ENVINF. **(a)** Relationship between number of DDCs, number of compute cores and relative speedup in comparison to a single core simulation (Color legend shows the value of relative speedup); **(b)** breakdown of the speedup curve AB (marked as dashed line in **a**) into speedup of calculation of chemical reaction i.e. IPhreeqc and other processes; **(c)** breakdown of the total time for chemical reactions, OGS#IPhreeqc interface and “other” for DDC = 4; **(d)** breakdown of the total time for DDC = 12.

A parallelization scheme to simulate reactive transport

W. He et al.

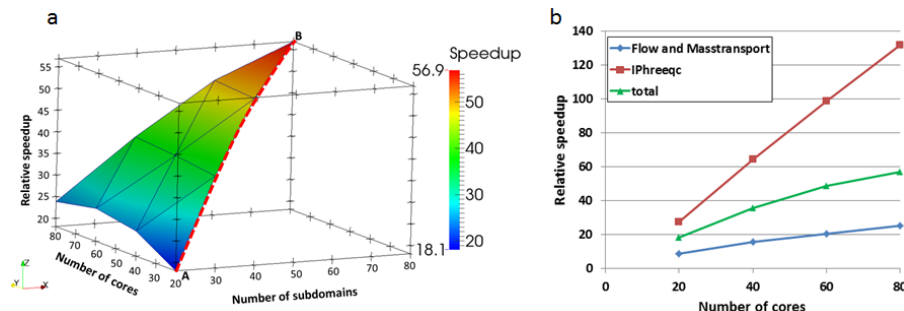


Figure 9. Performance of the parallelization scheme for the simulation of the 3-D test example on EVE cluster. **(a)** Relationship between number of DDCs, number of compute cores and relative speedup in comparison to a 2 compute cores simulation on ENVINF; **(b)** breakdown of the speedup curve AB (marked as dashed line in **a**) into speedup of calculation of chemical reaction i.e. IPhreeqc and other processes.

