PFLOTRAN: A massively parallel simulator for groundwater flow and transport

Richard Tran Mills[1], Peter C. Lichtner[2], Chuan Lu[2]
[1]National Center for Computational Sciences, Oak Ridge National Laboratory
[2]Earth and Environmental Sciences Division, Los Alamos National Laboratory

Abstract

Detailed modeling of reactive flows through geologic media is necessary to understand a number of environmental problems of national importance; examples are migration of radioactive fluids and geologic sequestration of CO2 in deep reservoirs. Such problems generally require simulations in three spatial dimensions and involve multiphase, multicomponent geochemical systems; typical simulations are very computationally demanding and might involve 10 or more chemical degrees of freedom on a grid of millions of nodes.

This poster describes PFLOTRAN, a massively parallel code for solving multiphase-multicomponent reactive flow and transport equations in nonisothermal, variably saturated porous media. PFLOTRAN is built on top of the PETSc parallel scientific toolkit and uses its efficient Newton-Krylov solver framework to solve the nonlinear equations arising from the fully-implicit timestepping scheme, using a domain-decomposition approach. We present performance timings from the MPP2 machine at EMSL/PNNL and the Cray XT3 machine (Jaguar) at NCCS/ORNL that demonstrate the extreme scalability of PFLOTRAN, and we illustrate some of science that is being done with the code.

Architecture of PFLOTRAN

Flow and transport modules

PFLOTRAN consists of two separate modules PFLOW and PTRAN that can be run either in standalone or coupled modes. PFLOW solves multiphase flow equations and PTRAN solves multicomponent reactive transport equations. In coupled mode, flow velocities, saturation, pressure and temperature fields computed from PFLOW are fed into PTRAN. For transient problems, sequential coupling allows changes in porosity and permeability due to chemical reactions to alter the flow field.

Modular design

Both codes employ a modular design and use Fortran 90 features to provide some degree of object orientation. Fortran 90 modules reduce unnecessary sharing of program variables and procedures. PFLOW uses modules and derived types to encapsulate data and the methods that operate on them. A PFLOW driver program loads `pflow_grid_module`, which provides a constructor function `pflowGrid_new` that returns a `pflowGrid` object. That object contains all data about the state of the variables on the flow grid, but the driver does not access those directly. Instead, method subroutines provided by `pflow_grid_module are used`, for example `pflowGrid_step` which advances the flow simulation one time step.

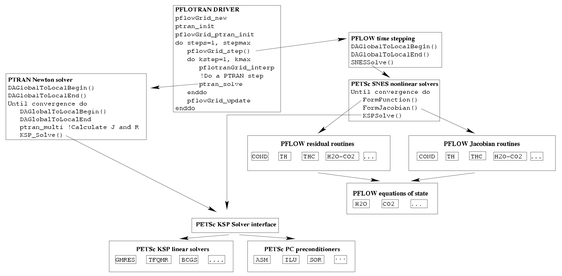


Figure: Cartoon representation of relationships between PFLOTRAN and PETSc components.

PETSc solvers

PETSc solvers (Balay et al., 1997) are used to solve the system of nonlinear equations arising at each timestep. PTRAN uses its own Newton solver routine, while PFLOW uses the SNES nonlinear solver framework of PETSc: this allows easy application of various trust region or line search techniques as well as methods for choosing the accuracy required for each Newton step. Jacobians can be evaluated analytically, or, if memory is at a premium, a matrix-free approach, in which the action of the Jacobian is calculated via finite differences, can be employed. We note that the modular implementation of the nonlinear solvers we use makes it convenient to add new physics to the code: a new residual function (and, optionally, a Jacobian calculation routine) for the Newton-Raphson equations simply need be added. Both codes use the PETSc `KSP` linear solvers and PC preconditioners to solve the linear systems arising at each Newton step; this allows the choice of any of several iterative methods (e.g. GMRES, TFQMR) or preconditioners at runtime. Parallel preconditioning has usually been accomplished using an additive Schwarz method with ILU applied to each subdomain. We are currently adding support for geometric multigrid techniques using the PETSc high-level `DMMG` multigrid framework.

Parallelism

Parallelism is achieved using a domain-decomposition approach. Each processor is assigned a subdomain of the system and a parallel solve of the system is implemented over all processors. Message passing is required at the boundary nodes to adjacent processors to compute flux terms (see Figure below). PETSc ``Distributed Array'' `DA` objects are used manage the distribution of field variables across the domains and message passing between domains. Ghost point scatters and gathers are handled conveniently by `DA` routines.

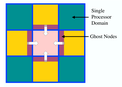


Figure: Schematic of domain-decomposition showing position of ghost nodes and message passing indicated by arrows.

Mathematical formulation

PFLOTRAN consists of two distinct modules: a mass and energy flow code (PFLOW) and a reactive transport code (PTRAN). The module PFLOW solves mass conservation equations for water and other fluids and an energy balance equation. The module PTRAN solves mass conservation equations for a multicomponent geochemical system. The reactions included in PTRAN involve aqueous species and minerals which can be written in the genereral form

$$\sum_j \nu_{ji} \mathcal{A}_j \rightleftharpoons \mathcal{A}_i,$$

and

$$\sum_j \nu_{jm} \mathcal{A}_j \rightleftharpoons \mathcal{M}_m,$$

respectively, where the set of species $\{A_j\}$ refer to a set of primary or basis species in terms of which all other species are written, A_j denotes an aqueous complex referred to as a secondary species, M_m refers to a mineral, and ν_{ji} and ν_{jm} are reaction stoichiometric coefficients derived from an extensive database.

Mass Conservation: Flow Equations

$$\frac{\partial}{\partial t}\phi s_\alpha \rho_\alpha X_i^\alpha + \nabla \cdot \left[\boldsymbol{q}_\alpha \rho_\alpha X_i^\alpha - \phi s_\alpha D_\alpha \rho_\alpha \nabla X_i^\alpha \right] = Q_i^\alpha$$

Energy Conservation Equation

$$\frac{\partial}{\partial t}\left[\phi \sum_\alpha s_\alpha \rho_\alpha U_\alpha + (1-\phi)\rho_r c_r T \right] + \nabla \cdot \left(\sum_\alpha \boldsymbol{q}_\alpha \rho_\alpha H_\alpha - \kappa \nabla T \right) = Q_e$$

Multicomponent Reactive Transport Equations

$$\frac{\partial}{\partial t}\left[\phi \sum_\alpha s_\alpha \Psi_j^\alpha \right] + \nabla \cdot \sum_\alpha \boldsymbol{\Omega}_j^\alpha = -\sum_m \nu_{jm} I_m + Q_j$$

Mineral Mass Transfer Equation

$$\frac{\partial \phi_m}{\partial t} = \bar{V}_m I_m$$

Darcy's Law

$$\boldsymbol{q}_\alpha = -\frac{k k_\alpha}{\mu_\alpha}\nabla(p_\alpha - W_\alpha \rho_\alpha g z)$$

Total Concentration

$$\Psi_j^\alpha = \delta_{\alpha l} C_j^\alpha + \sum_i \nu_{ji} C_i^\alpha$$

Total Solute Flux

$$\boldsymbol{\Omega}_j^\alpha = (-\tau\phi s_\alpha D_\alpha \nabla + \boldsymbol{q}_\alpha)\Psi_j^\alpha$$

Saturation

$$\sum_\alpha s_\alpha = 1$$

Mole Fraction

$$\sum_i X_i^\alpha = 1$$

Capillary Pressure

$$p_\alpha = p_\beta - p_{\alpha\beta},$$

Discretization

PFLOTRAN uses a finite-volume discretization on a regular grid. Time stepping is fully implicit (backward Euler). The scheme can be summarized as follows:

General Form of Partial Differential Equation

$$\frac{\partial A}{\partial t} + \nabla \cdot \boldsymbol{F} = Q$$

$$\boldsymbol{F} = \boldsymbol{q}\rho X - \phi D \rho \nabla X$$

Accumulation

$$\int_{V_n} \frac{\partial}{\partial t} A \, dV \simeq \frac{A_n^{t+\Delta t} - A_n^t}{\Delta t} V_n$$

Flux

$$\int_{V_n} \nabla \cdot \boldsymbol{F} \, dV = \int_{\partial V_n} \boldsymbol{F} \cdot d\boldsymbol{S} = \sum_{n'} F_{nn'} \mathcal{S}_{nn'}$$

$$F_{nn'} = q_{nn'} \rho_{nn'} X_{nn'} - (\phi D \rho)_{nn'} \frac{X_n - X_{n'}}{d_n + d_{n'}}$$

Source Term

$$\int_{V_n} Q \, dV \simeq Q_n V_n$$

Newton-Raphson Equations

$$\sum_{n'} J_{nn'}^k \delta x_{n'}^{k+1} = -R_n^k$$

Residual

$$R_n = \left(A_n^k - A_n^{k+1}\right)\frac{V_n}{\Delta t} + \sum_{n'} F_{nn'} \mathcal{S}_{nn'} - Q_n V_n$$

Jacobian

$$J_{nn'} = \frac{\partial R_n}{\partial x_{n'}}$$

Anomalous behavior of Pu at the Nevada Test Site

PFLOTRAN is being used to understand the anomalous migration of radionuclides at the Nevada Test Site, where 828 underground nuclear tests were conducted between 1951 and 1992.

The BENHAM test at the Nevada Test Site, with an announced official yield of 1.15 megatons produced a spherical cavity estimated at 200 m in diameter. The bottom of the cavity filled with melted rock, referred to as melt glass. A cylindrical, rubblized, chimney formed as rock above the cavity collapsed, extending from the working point of the test to above the TSA welded tuff aquifer. The melt glass is thought to contain a large proportion of the radionuclide inventory for Pu, whereas other less refractory radionuclides are distributed between the melt glass and the surrounding chimney material. 239,240Pu and other radionuclides from the BENHAM test have been observed in suprisingly distant sampling wells: one in a lava formation located near the depth of the working point (WP) and one in a TSA welded tuff aquifer located some 500 m above the WP situated at a depth of 1402 m. PFLOTRAN is being used to understand this behavior.

Massively parallel high performance computing is required to tackle these intrinsically 3D problems. Model simulations conducted using PFLOTRAN on PNNL/EMSL's MPP2 explain part of the puzzle by showing a pulse release of radionuclides to the TSA aquifer over a narrow window in time through the creation of convection cells caused by heat released from the melt glass (Lichtner and Wolfsberg, 2004).

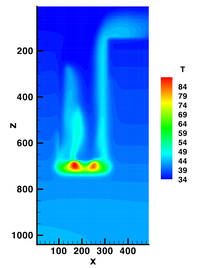


Figure: Cross section showing near-field temperature distribution after an elapsed time of 10 years.

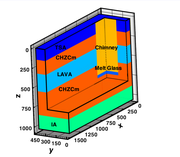


Figure: Cutaway view showing stratigraphy of grid in 3D simulations.

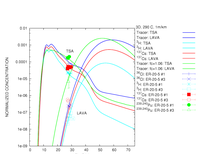
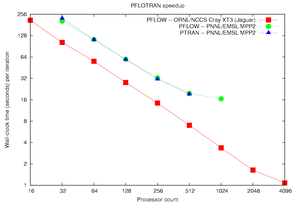


Figure: Breakthrough curves (amounts leaving boundaries) in the TSA and LAVA stratigraphic units.

Performance and scalability

PFLOTRAN been designed from scratch with parallel scalability in mind, and it displays excellent scaling characteristics on modern supercomputers. The figure at right shows the performance of PFLOW running a one phase thermo-hydrologic benchmark problem on a 256 x 64 x 256 grid with three degrees of freedom per node (approximately 12.6 million degrees of freedom total). The benchmark was run on both the MPP2 cluster at PNNL/EMSL, a cluster of 1960 1.5 GHz Itanium2 processors with Quadrics QsNetII interconnect, and Jaguar, the 5200 Opteron processor Cray XT3 at ORNL/NCCS. PFLOW scales quite well on both machines, bottoming out at around 1024 processors on MPP2, and scaling exceptionally well on Jaguar, displaying linear speedup all the way up to 2048 processors, and still displaying modest speedup when going from there to 4096 processors.

PTRAN scales similarly, which is not surprising because its computational structure is nearly identical to that of PFLOW. The figure shows the performance of PTRAN on MPP2 running a benchmark problem on 256 x 64 x 256 grid with four degrees of freedom per node.



CO_2 sequestration: Density driven fingering

A possible mechanism for mitigating the effects of CO_2 on global climate change is injection of supercritical CO_2 into deep underground reservoirs to prevent its release to the atmosphere. PFLOTRAN is being used to assess CO_2 sequestration in geologic formations and to evaluate the dissipation of a CO_2 plume as it evolves over time. The model simulations will also be useful in assessing leak rates to the surface, for example, through abandoned wells and fractures in cap rocks. Accurate assessment of the retention characteristics of candidate subsurface geologic formations is crucial: a poor choice of formation will defeat the goal of sequestration when CO_2 escapes back into the atmosphere. Additionally, sudden release of sequested CO_2 can have catastrophic consequences.

Characteristic of this problem is the formation of density-driven instabilities resulting from an increase in fluid density as CO_2 dissolves into H_2O, and viscous fingering caused by reduced viscosity of CO_2 compared to H_2O. The figures depict CO_2 fingering after 200 simulation years and the resultant mineral alteration. Mineral alteration is minimal, though it is expected that given longer times (~10^4 years), significantly greater alteration will occur. (Lu and Lichtner, 2005)

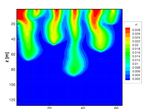


Figure: Mole fraction of CO_2.

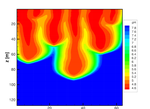


Figure: pH.

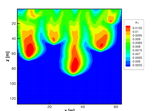


Figure:K^+.

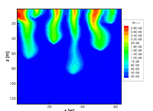


Figure: Al^{3+}.

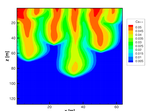


Figure: Ca^{2+}.

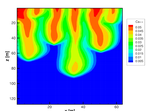


Figure: $SiO_{2(aq)}$.

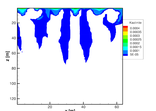


Figure: Kaolinite

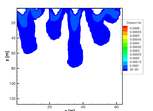


Figure:Dawsonite.

References

Balay, S., Eijkhout, V., Gropp, W. D., McInnes, L. C., and Smith, B. C. 1997. Efficient Management of Parallelism in Object Oriented Numerical Software Libraries. Modern Software Tools in Scientific Computing. E. Arge, A. M. Bruaset and H. P. Langtangen, Ed., Birkhauser Press, Boston, MA, 164-202.

Lichtner, P. C. and A. Wolfsberg, 2004. Modeling Thermal-Hydrological-Chemical (THC) Coupled Processes with Applications to Underground Nuclear Tests at the Nevada Test Site: A Grand Challenge Supercomputing Problem. MPU Workshop: Conceptual Model Development for Subsurface Reactive Transport Modeling of Inorganic Contaminants, Radionuclides and Nutrients, Los Alamos National Laboratory, Report: LA-UR-04-1826.

Lu, C. and Lichtner, P. C., 2005. PFLOTRAN: Massively Parallel 3D Simulator for CO2 Sequestration in Geologic Media. Fourth Annual Conference on Carbon Capture and Sequestration, DOE/NETL.

Acknowledgements

One of the authors was supported by a Department of Energy Computational Science Graduate Fellowship, administered by Krell Institute, while performing portions of this work. This research was performed in part using the Molecular Science Computing Facility (MSCF) in the William R. Wiley Environmental Molecular Sciences Laboratory, a national scientific user facility sponsored by the U.S. Department of Energy's Office of Biological and Environmental Research and Located at Pacific Northwest National Laboratory (PNNL). PNNL is operated for the Department of Energy by Battelle. Some of the timings presented were obtained using resources of the National Center for Computational Sciences at Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract no. DE-AC05-00OR22725.