Title: Parametric Visualization and Computation of Large Geochemical Datasets

Principal Investigators: W. Bethel, NERSC/ICSD; Karsten Pruess, ESD; George Brimhall, UC-Berkeley.

1.0 Purpose/Goals

This project seeks to make advances in geochemical modeling tools by using computational facilities at Berkeley Lab; seeks to make advances in techniques for visualization of large datasets such as those produced by geochemical simulation models on Berkeley Lab/NERSC equipment; and seeks to make advances in techniques for visualization of divergent data from different sources, some of which is computed and some of which is observed.

This project is motivated by three orthogonal but related goals:

- The need for advances in visualization tools which can process very large amounts of data.
- Implementation and interactive visualization of Brimhall's "First-Principle" modeling tool on Berkeley Lab/NERSC facilities, and visualization of results. (need better words?)
- The need for integration, via visualization, of heterogeneous data, especially the results of numerical models run on NERSC machines.

Existing tools for visualizing data are inadequate for use on "large" datasets, such as those produced by models by Berkeley Lab researchers. As computing capacity increases, researchers scale up computational tasks in size, thereby producing results that are one- to two-orders of magnitude larger in size than those used in recent years. Existing visualization tools, those that attempt to visualize and display all of the data at once, are impractical. A primary goal of this project is to systematically develop techniques for visualization of large data sets using hierarchical and multi-resolution methods.

Model verification is an important goal for researchers. A forward "first-principle" modeling project is currently underway coupling ground water fluid flow and chemical kinetics to simulate district-scale transport processes controlled by global climate change as a function of time. However, due to the computational constraints imposed by a desktop environment, interactive manipulation and computation of this model is not possible. This project seeks to move this model to an MP machine at NERSC, and to couple it with existing and new visualization tools in the visualization laboratory.

In a related LDRD (Pruess/Brimhall: Reactive Chemical Transport in Geologic Media), there is a large amount of data from various sources which has been partially integrated into a single display and is the subject of a demonstration in the Visualization Laboratory. The first-principle model (above) will produce data which will be integrated into the existing model. Additionally, there is new data coming from the El Salvador Mining District in northern Chile which the Visualization Group has been asked to also integrate into the model, including simulation data produced by Pruess at NERSC facilities. The work described in this proposal will be closely coordinated with the work in the Reactive Chemical Transport LDRD.

2.0 Approach/Methods

2.1 Visualization of Large Data (Bethel)

When visualizing large datasets, two orthogonal sets of issues must be balanced. One issue is the fact that many large datasets simply cannot fit onto the visualization machine. A separate, but related issue is the identification of a visualization technique for a large dataset.

With existing tools for visualization, the entire data set is used in creating images. At best, the image often appears very cluttered, or noisy. Deriving meaning or understanding from such images is not possible. The themes present in successful visualization tools for large data sets will include:

- Multiresolution representation of the data itself so that existing visualization tools can be used on only a portion of the large data.
- Multiresolution visualization such that a "region of interest," specified by the user, contains a higher level of visual detail.

One focus of this project is on ways of specifying the "region of interest." Existing tools in the Visualization Laboratory include an infrastructure for use of six-dimensional input devices. These tools have been used successfully in a variety of applications (Bethel, et. al, 1994; Bethel, 1996). New tools for specifying regions of interest can be built upon this existing infrastructure. For example, one way for making the complex, three-dimensional data more "intelligible" is to define a two-dimensional virtual slice plane that "cuts" through the three-dimensional volume, and display a suitably interpolated data field (e.g., of ore grades) in that plane, then make the whole plane "dynamic" by allowing the user to "drive" it at will through the three-dimensional volume. For orientation, a gnomon should be present "on the side" to show the instantaneous orientation of the plane in the volume. To aid in comprehension of three dimensional structure (Ware, 1996) we will use a stereo-augmented display system.

A second focus of this project is to identify ways in which the data may be represented hierarchically such that the full model need not reside entirely in memory for the purposes of visualization. For example, high-level-of-detail portions of the dataset which are "behind" the user can be held in secondary storage, and transferred to primary storage at such time that those blocks of data lie within the user's field of view. Existing work has been done to represent volumetric data using a hierarchical representation, with error or tolerance computed at each level in the tree [Laur and Hanrahan, 1994]. The use of wavelet transformations, as are currently applied in image processing, hold much promise an alternate and more compact method for hierarchical representation.

2.2 First-Principle Modeling Tool (Brimhall, Bethel)

Empirical parametric equations (Brimhall and Alpers, 1985; Alpers and Brimhall, 1989) based solely on mass balance can be used to calculate the 3-dimensional changes in elemental concentrations during the time-integrated history of an ore deposit caused by the falling ground water table. These mass conservation expressions relate the metal concentrations in the three sub-systems created by conditions above and below the ground water table. The uppermost subsystem is the zone above the ground water table which is progressively leached of its metal by oxidation. The intermediate zone is directly beneath the leached zone and represents the enriched state when metal leached from above has be carried downward by migrating ground water and reprecipitated under reducing conditions. The lowest zone is the parent material unaffected by the flow of ground water.

2.2.1 Parameters

The concentration of copper in the leached zone is *l*, that of the enrichment zone is *b*; and that of the proto-ore is *p*. The thickness of the leached zone is *Lt* and the thickness of the enrichment zone is *B*.

Omitting rock density terms for the sake of simplicity here, the mass conservation equation expressing *b* as a function of the other terms is $b = p + (L_t/B)(p-l) + \text{Lateral } \dot{\text{F}}\text{lux}/B$

2.2.2 Implementation

From the database in hand on the El Salvador district in Chile, we can calculate the parametric equations in three dimensions. Hence, we can create a purely empirical simulation of the entire district in terms of its chemical evolution. In other words, we know what the initial conditions were and we also know the final distribution of metal concentrations. This parametric approach allows us to calculate progressive increments in mass concentration and thereby output successive intermediate stages which can then be animated in rapid succession to create a movie of the overall transport field including the regions near permeable faults where lateral instead of vertical fluid flow has occurred.

2.3 Heterogeneous Data (Bethel, Pruess, Brimhall)

We will combine various species of data into an integrated display. The types of data to be integrated include topographic data, assay data, physical locations of structures such as well bores, geological information such as fractures and lithology, computed data including the results of the parametric geochemical model, computed hydrodynamic data, and so forth.

Multiresolution issues are pervasive in this; e.g., water flow and mineral redistribution may be dominated by paleochannels, fault systems, and lithologic units on the large scale, and a hierarchy of fractures on progressively smaller scales. "Complex" coupled physical and chemical processes are being played out in "complex" geologic settings, eventually producing the kinds of ore grade distributions observed today. Visualization efforts start by addressing the final outcome, e.g. the distribution of ore today, then go after configurations and patterns in the geologic past, first in a "static" way (snapshots in time), eventually dynamically (as "movies").

As part of the effort, we wish to port the Khoros scientific computing environment onto the T3E at NERSC. Khoros supports MPI and comes equipped with an array of computational and visualization tools. Some of these tools are already in use in the Visualization Laboratory. The company which produces the Khoros software is interested in assisting LBL with this project, and will provide gratis support.

3.0 Potential Results or Significance

Implementation of the described first-principle model of water flow and chemical kinetics on NERSC facilities and coupling it with interactive visualization tools will provide a powerful and compelling environment for validity checking and further model refinement.

The problems associated with the visualization of large and heterogeneous data transfer to other disciplines and applications, such as detector modeling, study of nuclear storage facilities, and so forth.

At present, there is no well-established path between the computational facilities at Berkeley Lab/NERSC and three dimensional visualization workstations. The work outline in this proposal implements such a path.

The results of this project can be encapsulated into one or more demonstrations which can be copied to the machine in the Washington DC office. In addition, the effort described in this project will facilitate similar demonstrations of other projects, such as the Yucca Mountain Storage Facility project.

4.0 Relationship to Other Berkeley Lab Projects

"Reactive Chemical Transport in Geologic Media" (Pruess/Brimhall).

5.0 Proposed Organization

Bethel - Visualization and software infrastructure on NERSC MP facilities; Brimhall and GSRA- Parametric Geochemical model implementation; Pruess and RA- Hydrodynamic code extinctions to support visualization.

6.0 Budget

$150,000 effort for FY98 distributed as follows:

6.1 Effort

- LBL (NERSC/ICSD); 1/2 FTE (Bethel) and 1/2 FTE (Postdoc or CSE II equivalent) - $60K + $60K.
- UC-Berkeley: 1/4 FTE (0.08 FTE for Brimhall, 0.17 FTE for GSRA from Geology department to do programming.) - $15K.
- LBL (ESD); 1/4 FTE (0.08 FTE for Pruess, 0.17 for Research Assistant to do programming.) - $15K.

6.2 Facilities

- Disk space and CPU time on the NERSC T3E (1 gigabyte, 50 hours).

7.0 References

Brimhall, G. H and Alpers, C. A., 1987, Analysis of supergene ore-forming processes and ground water solute transport using mass balance principles: Econ. Geol., v. 80, p. 1227-1256.

Alpers, C. A and Brimhall, G. H, 1989, Paleohydrologic evolution and geochemical dynamics of cumulative supergene metal enrichment at La Escondida, Atacama Desert, northern Chile: Econ. Geol., v. 84, p. 229-255.

Ware, C. and Franck, G, "Evaluating Stereo and Motion Cues for Visualizing Information Nets in Three Dimensions," ACM Transactions on Graphics, Volume 15, Number 12, April 1996.

Bethel, W., Jacobsen, J., and Holland, P., "Site Remediation in a Virtual Environment," LBL-35262/UC-405, Proceedings of IS&T/SPIE Symposium on Electronic Imaging Science and Technology, San Jose, California, January 1994.

Bethel, W, "Implementing Virtual Reality Interfaces for the Geosciences," LBL-38618/UC-405, Proceedings of Virtual Reality in the Geosciences, Halden, Norway, June 1996.

Laur, D., and Hanrahan, P., "Hierarchical Splatting: A Progressive Refinement Algorithm for Volume Rendering," Computer Graphics 25,4, Proceedings of Siggraph 1991.