Swat Team Trip Report

University of California at Los Angeles
Los Angeles, CA
Monday, 2 November 1992

Mike Olson (mike_olson@postgres.Berkeley.EDU)

1. Introduction

On Monday, 2 November 1992, the Sequoia 2000 Swat Team visited the University of California at Los Angeles. The meeting included swat team members (Jim Frew, Mike Olson, Peter Kochevar, Duane Waliser, and Allison Woodruff), Department of Water Resources meteorologist Frank Gerhke, and researchers and staff members at UCLA (Fabrice Cuq, Michael Ghil, Roberto Mechoso, David Neelin, Richard Turco, Bill Weibel, Ruiyan Lin, and Kayo Ide).

As was the case at CRSEO in Santa Barbara, researchers gave short presentations of their projects, and the group discussed visualization, data lineage, and computational steering.

2. Swat Team Trip Reports

Both Michael Ghil and Roberto Mechoso expressed concern over the detailed benchmarks described in previous swat team trip reports. For the very near term — before the January retreat — the benchmarks described are far too ambitious for implementation.

UCLA researchers would like to focus on two benchmarks before January.

First, it should be possible to combine and display two or three two-dimensional fields. For example, satellite images and field observations should be able to be displayed on a map. There should be different ways to do this visualiztion — superimposition of the images on the same canvas, displaying the datasets side-by-side on the screen, or laying one dataset on top of the other, with space between them, and letting the user see the bottom dataset by displaying the pair at an angle.

Second, it should be possible to manage and visualize four-dimensional data, and specifically spatial data over time. Slices, projections, and animation should be supported.

Olson stated that the swat team trip reports were intended to publicize the long-term goals of the project, and were not a commitment to implement all of the services they describe by January. A document describing the short- and longer-term goals of the project with respect to visualization, data lineage, and computational steering will be written after the swat team has visited all of the sites. This document will be available before 22 November 1992.

3. Visualization

In addition to the two short-term services described above, climatologists at UCLA would like other services from the data management and visualization software. This section gives a brief description of the major services desired.

It would be useful to have simple generic processing tools built in, such as bandpass filters and image operators. Image operators implement, for example, subtraction of images — pixels in one image are subtracted from pixels in another, producing a new image.

Intercomparison of datasets (such as displaying them on the same map) requires a tool to do projection conversion. All four of the sites visited have now requested such a tool.

UCLA researchers would prefer not to be required to use a single technique for doing space-variant data interpolation, such as optimal interpolation, since space-variant interpolation is an area of active research in the Earth science community. Any interpolation tool must be extensible.

Richard Turco urged the group to come up with an icon-based, Mac-like visualization tool that is easy for the scientists to use without consulting a computer programmer.

Any visualization tool should be able to produce images for use in conference presentations and in publications. The best interface would be for the scientist to be able to press a button to get a publishable figure from the screen image, and another to get a slide for a talk. The tool should permit citations and titles to be appended. Some default output format for these other media — black and white contour plots standing in for color maps, for example — should be provided.

Finally, researchers at UCLA were concerned with how they could get experience with the tools available under Sequoia. Peter Kochevar will return to Los Angeles for a one-day visualization session on 24 November, and may bring other members of his group from the San Diego Supercomputer Center.

At the November 24 meeting, Peter will use AVS to visualize some of Mechoso's GCM data. Both Peter and Frew got pamphlets describing the format of the GCM data. Peter's work with John Roads at Scripps is intended to give the project experience with visualization and an initial working system, which can be modified and extended for more general use.

4. Data Lineage and Computational Steering

UCLA researchers agree that the services described in the CRSEO trip report are interesting. They caution that for the short term, the project should concentrate on simple visualization benchmarks.

5. Data Management Issues

Duane Waliser has loaded a good deal of Coastal Zone Color Scanner (CZCS) data into Bigfoot under the Inversion file system. Very little tabular data of interest to people at UCLA has been loaded into POSTGRES to date.

Two important issues with respect to data management are standardization of formats and how and where to load data. Jim Frew chairs the sequoia-schema working group to deal with formats and metadata management, and researchers at all sites are encouraged to participate in this effort. Selection of format must be driven by Earth scientists, not computer scientists. Frew is also able to allocate space for loading of data.

At the 24 November meeting with Peter Kochevar and his group, UCLA researchers will discuss what they need in terms of data management and data loading.

5.1. Formats, Operators, and Data Loading

Roberto Mechoso asked for guidance on what should be done with data that is currently ready to load onto Bigfoot. For example, extant GCM output and observational data are in different projections and are at different resolutions. The Earth scientists are willing to convert data formats to whatever is in use in the project, but they do not want to have to do this conversion more than once.

Frew gave the following response:

- In the short term, researchers should treat the Bigfoot data store as a big file system. Data should be loaded in whatever format is currently available. The intent is to permit people to use the storage system to do work immediately.
- In the medium term, the computer scientists, in cooperation with the Earth scientists, will consider what operators and programs should be available to manage this data, and will add these to POSTGRES for use at all sites. Also, a standard format will be selected, and tools to do conversion automatically will be constructed.

Selecting a representation for scientific data is inherently an iterative process — formats will be selected and refined as their shortcomings are identified.

5.2. New Hardware

Mechoso and Michael Stonebraker were jointly awarded a grant by NASA to study high-performance computing systems. Their goal is to implement a parallel distributed global climate model on a network of Alpha workstations. This distributed computer would like to be a Bigfoot client, and any storage management protocol should consider it. The system will do modelling and visualization interactively, and will produce a timestep image every ten seconds of wallclock time.

Peter Kochevar and Jim Frew agreed to help Stonebraker and Mechoso draft a proposal for DEC requesting an equipment grant for Alphas to do the HPCC project.

6. Projects at UCLA

This section gives a very brief description of some of the major projects underway at UCLA.

6.1. Exploring Model Parameter Spaces

Both Michael Ghil and David Neelin are trying to explore the parameter space for complex models. The strategy is the following:

- Using some knowledge about physics, climatology, or a model's implementation, figure out what parameter variations are likely to be interesting. For example, if you change the solar constant or the length of the day, will the model do interesting things?
- Run a global climate model to verify that there's actually something interesting happening when you make the changes described. This requires computational steering and interactive visualization.
- After verifying that the GCM behaves as expected, try to analyze the behavior on a simpler climate model. Smaller models are easier to solve analytically, and so can give clearer understanding of basic physical processes than a complex global model.

Varying "constants" in this way is not as strange an idea as it sounds at first to computer scientists. For example, the rate of rotation of the Earth actually varies with time. The planet's rotation imparts kinetic energy to its atmosphere (the mountains push the air), so the planet slows down. Later, the air runs into other mountains, returning the energy to the earth. The variation in the length of the day can be on the order of milliseconds — a surprisingly long time, when you consider the inertial forces involved.

6.2. Visualizing Chemical Tracer Trajectories

Richard Turco and his students are working on the visualization of atmospheric flows. Model data are displayed over satellite images, and tracers are inserted in the image. The model output is then animated, and the paths of the tracers are displayed. At present, they're able to put a small number of tracers (fewer than ten) in an image. Eventually, they would like to do thousands.

The tracers are not simple points; they are volumes of air with internal state (for example, pressure and temperature). The visualization tool does not currently display this state, and they would like some guidance on how it might do so.

This is the first extensive visualization using AVS that the swat team has seen that was not done by Peter Kochevar. The person who did the AVS work at UCLA is Ruiyan Lin (ruiyan_lin@postgres.CS.Berkeley.EDU); climatologists interested in using AVS at other sites may be able to get some guidance from him.

6.3. Visualizing Atmospheric Flows

Christopher Strong and Michael Ghil are working on the visualization of atmospheric oscillation patterns over the Northern Hemisphere midlatitudes. The period of these oscillations is typically between 30 and 60 days. The data are obtained from a one-layer model and can therefore be animated by way of a 2-D sequence of images. The visualization package being used is Interactive Data Language (IDL); they like this package because of its ability to filter model data, perform power spectrum analysis, and compute statistics for a wide range of model or observational data. Finally, its animation module is easy to apply, and provides an advantage over traditional snapshots of model fields, allowing them to view the evolution of a selected pattern in real time.

In January at the Lake Arrowhead retreat, Strong and Ghil plan to present a three-way intercomparison of these 30- to 60-day oscillation patterns through a hierarchy of data extracted from

- the simple, one-layer model;
- the UCLA General Circulation Model; and
- real atmospheric observations.

Using visualization tools such as IDL, they plan to illustrate how animation is a powerful interactive tool in discerning the behavior of numerical models compared with that of the real atmosphere.

7. Miscellaneous

This section covers a side issues that arose during the day.

7.1. Neural Networks

NASA is funding another project at UCLA to recognize visual patterns in images. For example, one application would be to find Hurricane Andrew in a satellite picture of the eastern coast of North America. This would lend itself to content-based indexing of images.

Roberto Mechoso will see whether Sequoia can get the results of, or get involved in, this effort. A researcher at SDSC is looking for a good subject for a neural network application, and this may fit the bill. Kochevar will put this researcher in contact with Mechoso.

7.2. Goals of the Swat Team

An explicit goal of these swat team trip reports is to give researchers at all sites a detailed idea of who is active at other sites, and what projects they are working on. In particular, names and areas of expertise are included so that researchers can contact one another directly.

Project members are strongly encouraged to use these reports to get to know one another, and to develop ties with peers at other sites.

The final swat team report will describe the visualization benchmarks we will support in the short and long term. The trip reports will be included in that document as appendices.

7.3. Hardware Problems

UCLA purchased two input devices for their Decstation 5000s to be used with AVS. The devices consist of several knobs and a panel of pushbuttons. AVS on most platforms permits users to control visualizations using such an input device.

It turns out that AVS on the Decstations does not support this input device. Researchers at UCLA are disappointed. Frew will talk to Ira Machefsky and Peter Kochevar about this, so that both DEC and AVS are aware of the problem. DEC personnel with questions on the issue should contact Fabrice Cuq (fabrice_cuq@postgres.CS.Berkeley.EDU).

It is important that these devices be supported for visualization.

7.4. Alphas

Researchers at UCLA would like a FORTRAN compiler for the Alpha.

7.5. POSTGRES Client Software

Mike Olson will make a distribution package available to Sequoia sites that contains only the client software for POSTGRES. Sites need not copy and compile the full POSTGRES 4.0.1 distribution in order to use POSTGRES on heel.

8. Summary

The trip to UCLA resulted in a finer focus on short-term goals for visualization. Issues in data lineage and computational steering, while of interest to scientists there, were not considered as urgent as getting a good initial visualization tool.