Visualisation of large datasets with Paraview

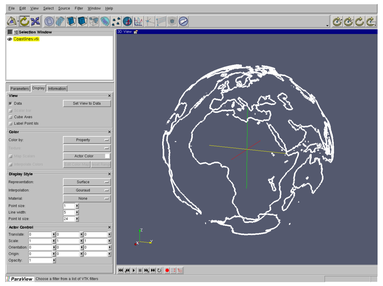
Christoph Moder, Hans-Peter Bunge, Heiner Igel, Bernhard Schuberth
Department of Earth and Environmental Sciences (Geophysics), Ludwig-Maximilians-Universität München
Contact: *mail@christoph-moder.de*

Abstract

With growing computing power, simulations become more sophisticated and higher resolutions can be handled. Together with the increasing use of parallel computers, this leads to large amounts of data. This data explosion is posing mounting difficulties in the post-processing and visualisation phase of modern simulations. Whereas large-scale simulations nowadays can often be split up on parallel computers, standard visualisation methods still tend to operate under the traditional sequential computing paradigm and hence the visualisation process must deal with the entire dataset at once. Thus often it does not fit even on quite large PCs. In our investigations, we found that the new open-source parallel visualisation program Paraview instead is quite a good tool — it has a very flexible architecture allowing it to run on a computing cluster, so that even large datasets can be handled. Additionally, interactive processing of the data, hardware-acceleration and optionally 3D on a Geowall are supported. This makes it suitable for processing large-scale geophysical simulations like high-resolution geodynamic or seismic simulations with isosurfaces, cross sections and topography.

What are Paraview and VTK?

- VTK is an object-oriented high-level library for data processing and visualisation.
- Basic concept: a pipeline from the data source to the data sink via an arbitrary number of filters
- The library is written in C++, but there are bindings for scripting languages like Tcl and Python.
- Hardware acceleration via OpenGL is supported.
- Paraview is a visualisation program based on the VTK library.
- Both packages are available as open source software.

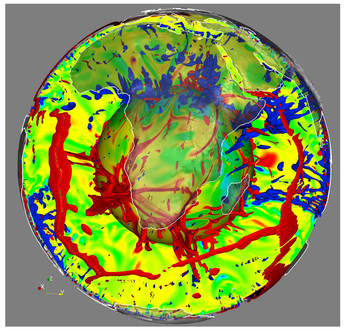


Paraview GUI

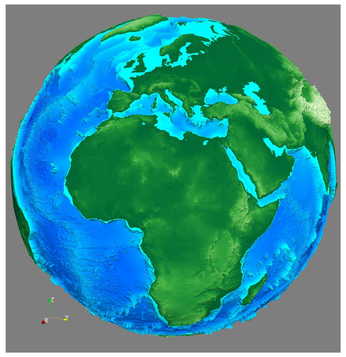
Cluster

- VTK and Paraview can be linked against an MPI library, so they can be used on computing clusters.
- Multithreading is supported, which is beneficial e.g. on multi-core architectures.
- Paraview implements a further abstraction layer for the separation between data server (stores the data and computes the filter operations), render server (computes the polygonal representation) and client (provides the GUI).
- This allows a very flexible setup. Large datasets can be processed if the data server is running on several machines, whereas the rest can be done on a single PC. This flexibility was one of Paraview's design goals — every available hardware should be used as effectively as possible.

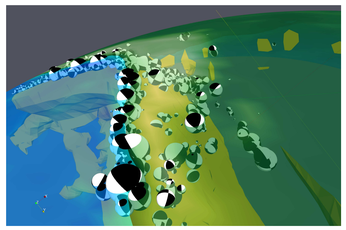
Examples



Simulation with a viscosity gradient, isosurfaces at $dT = -500$ K and +300 K



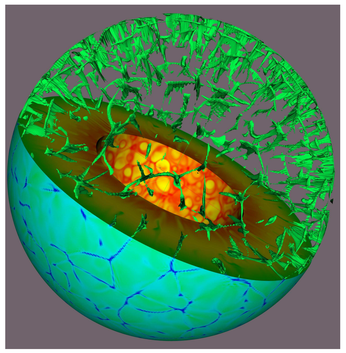
ETOPO dataset, 50 times exaggerated



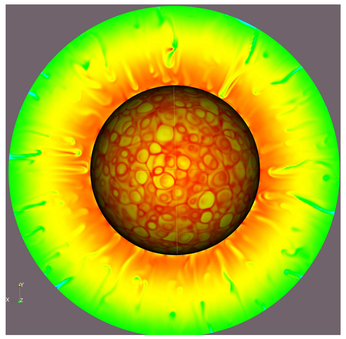
Earthquakes and subducting slab in South America

```
#!/home/SOFTWARE/VTK/bin/tclsh8.4
package require vtk

vtkPolyDataReader reader
        reader SetFileName [lindex
          $argv 0]

vtkPolyDataWriter writer
        writer SetInput [reader
          GetOutput]
        writer SetFileName [lindex
          $argv 1]
        writer SetFileTypeToBinary
        writer Write
exit
```

Tcl script for VTK file conversion from ASCII to binary



Convection in a isoviscous mantle, isosurface at 800 K



Convection in a isoviscous mantle

3D Stereo Mode

- VTK supports various stereo modes, including one suited for LCD shutter glasses.
- Some dual head graphics cards (e.g. Nvidia Quadro series) can distribute images with stereo information for shutter glasses on the two monitors.
- A *Geowall* is a room with two digital projectors (using DLP technology), a metallic screen and polarising filters in front of the projectors. Together with a PC with such a dual head graphics card, this yields a low-cost 3D projection facility, compared to a CAVE setup.
- Advantages: The audience needs only cheap polarising glasses ("passive stereo"), compared to LCD shutter glasses. 3D is possible in full color, unlike to stereo modes requiring e.g. red-blue glasses.
- Paraview can be patched to use this 3D mode; only one line in the source code has to be changed.

Data Processing

- The VTK files can be written in ASCII, thus formatted output from Fortran code can easily be used, and some low-level preprocessing (e.g. coordinate transformation) can be done with Unix shell tools like *awk*.
- Parallel file formats are available, which link together several subfiles. Since every simulation process on the cluster writes its data independently, these files can be converted into such subfiles without the necessity to assemble them into one big data file.
- The conversion to binary formats, as well as high-level preprocessing (e.g. isosurfaces), can be done with VTK scripts (see example on the left).
- The result can be viewed with Paraview — either running on the cluster or, if the amount of data has been sufficiently reduced (e.g. with isosurface filters), on a single PC.

Conclusions

- Paraview offers a fast and convenient way to view simulation data and do processing like isosurfaces or cutting.
- Coastlines and topography can easily be added, as well as further datasets.
- With a set of VTK scripts and shell scripts the preprocessing of the data can be automated very well.
- The software can be used on a single PC (32 bit version) as well as on a computing cluster (64 bit version); this allows to handle large datasets (more than 80 million grid points).
- With a visualisation cluster (equipped with fast 3D graphics cards) as frontend, the performance could be further improved.