SAND2011-2579P

Large Scale Visualization with ParaView

Abstract

ParaView is a powerful open-source turnkey application for analyzing and visualizing large data sets in parallel. Designed to be configurable, extendible, and scalable, ParaView is built upon the Visualization Toolkit (VTK) to allow rapid deployment of visualization components. This tutorial presents the architecture of ParaView and the fundamentals of parallel visualization. Attendees will learn the basics of using ParaView for scientific visualization with hands-on lessons. The tutorial features detailed guidance in visualizing the massive simulations run on today's supercomputers and an introduction to scripting and extending ParaView. Attendees should bring laptops to install ParaView and follow along with the demonstrations.

Detailed Description

General description and tutorial goals

We are proposing a half-day tutorial on ParaView for parallel/distributed visualization of large scientific data sets. This tutorial is going to present the distributed/parallel visualization capabilities of the open-source turn-key application ParaView and, by proxy, the Visualization Toolkit (VTK) upon which ParaView is based.

ParaView is designed from the ground up to run efficiently on large parallel distributed-memory clusters computers, and our current usage of ParaView has proven its success in this regard. In this tutorial we will dedicate time to focus on the features of performing large-scale parallel visualizations and scripting.

Targeted audience, content level, and relevance to SC attendees

We expect the content break down to be: 60% beginner, 40% intermediate, 0% advanced.

This tutorial has a wide appeal for all levels and types of attendees. Attendees familiar with visualization, graphics, or VTK will learn about the design an implementation of a successful end user application. Others will learn how to use ParaView to process their large data sets.

Beginners will benefit by learning how to configure and run ParaView and process their large data set in parallel. They will learn how the advanced processing modules operate internally. This understanding will allow them to select which visualization modules should be used for specific tasks.

Intermediate users will be given a guided tour on the inner workings of ParaView. Special attention is placed on establishing a proper platform and launching the application in a parallel environment. We also identify the pitfalls inherent in visualizing memory straining data. With an introduction to the Python scripting, we demonstrate how to customize and automate the ParaView application for domain-specific visualization solutions.

The majority of the Supercomputing attendees are from universities or government labs. In this environment it is important to be able share tools and applications. Since this suite of tools is open source, there are no barriers to collaboration between diverse organizations. The fact that there are no license fees for these applications is also important. Some university research groups may not be able to easily purchase expensive licenses for proprietary visualization applications.

Also, ParaView is a world leader in high performance visualization on distributed clusters and leadership-class supercomputers. Many Supercomputing attendees are researching the application of large parallel computing toward high performance computation and visualization. ParaView has been designed from inception to run well on distributed computing platforms.

Visualization applications are important tools for debugging simulations, sharing results and marketing supercomputing research. Knowledge of these flexible and configurable tools will benefit a wide audience of conference attendees (not only ones specializing in visualization).

Audience prerequisites

We have no expectations of the audience although an understanding of visualization techniques will be helpful.

Ensuring cohesive content

The presenters are divided by function, not content. One presenter will function as an instructor, providing the verbal content and leading the class. Another presenter will function as a demonstrator, showing the material in action while it is being discussed. The remainder of the presenters acts as tutors to give one-on-one help to participants. As one presenter controls the content of the tutorial, the content remains naturally cohesive.

Tutorial updates for SC

The material for this tutorial will be similar to the ParaView tutorial we gave in several other supercomputing conferences (SC07 – SC09). The attendance for these previous tutorials was good and the evaluations gave very positive feedback. As before, many of the presenters will function as tutors for one-on-one help with the attendees when they need it. As always, we will update the material for the latest features of ParaView. We will also grow our instructions on scripting, which has both become easier and is growing in importance for large-scale visualization.

We are also submitting a more advanced companion tutorial on in-situ visualization for those needing programming control over large-scale visualization.

Acknowledgements

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Description of Demonstration

The information in this tutorial is communicated via slide presentations and live demonstrations, both from a laptop provided by the presenter. We do *not* intend to provide training computers for the attendees, but the attendees should bring their own laptop to follow along with the tutorial.

Required Hardware

Our tutorial will be most effective if we can show both information slides and a live demonstration at the same time. For this, we will require **two projectors** and screens. Live demonstrations will be performed on a laptop provided by the presenters. We would also like a **wired Ethernet connection** to the Internet to demonstrate ParaView's remote visualization capabilities.

Attendees who wish to follow along with demonstrations will need to bring their own laptops. However, we also require wireless networking in the room to help them set up their computers with ParaView.

Outline

1. Introduction/overview/installation (15 minutes).
2. Basic Examples (1 hour, 30 minutes).
    a. User Interface.
    b. Sources.
    c. Loading data.
    d. Filters.
    e. Multiview.
    f. Volume rendering.
    g. Plotting.
    h. Time.
    i. Annotation.
    j. Animations.
3. Visualizing large models (45 minutes).
    a. ParaView architecture (client/server).
    b. Launching client/server.
    c. Rendering.
        i. Basic parameter settings.
        ii. Basic parallel rendering.
        iii. Parallel render parameters.
        iv. Parameters for Large Data.
    d. Parallel visualization algorithms.
        i. Ghost levels.
        ii. Data distribution.
    e. D3 Filter.
    f. Matching job size to data size.
    g. Avoiding data explosion.
    h. Culling data (threshold, clip, etc.).
4. Scripting (30 minutes).
    a. Python interpreter, servermanager module.
    b. State file saving/loading.
    c. Programmable filters.

Tutorial Notes Release

The presenters of this tutorial agree to release the tutorial notes on the SC11 USB stick.

Travel Support Request

Some of the presenters will request support for travel.

Résumé: Kenneth Moreland

Education

- Doctor of Philosophy Computer Science, University of New Mexico, Albuquerque, NM, July 2004.
- Master of Science Computer Science, University of New Mexico, Albuquerque, NM, May 2000.
- Bachelor of Science Computer Science, New Mexico Institute of Mining and Technology, Socorro, NM, May 1997.
- Bachelor of Science Electrical Engineering, New Mexico Institute of Mining and Technology, Socorro, NM, May 1997.

Work Experience

ParaView 3 Development Lead — Sandia National Laboratories, 10/2006–Present

Lead the ASC funded development effort for ParaView 3, a large-scale general visualization solution.

Scalable Visualization — Sandia National Labs, 8/2000–Present

Researched and developed parallel rendering codes and other scalable visualization algorithms targeted at performing scientific visualization on cluster computers.

Product Realization Environment — Sandia National Labs, 5/1996–8/2000

Developed and deployed a CORBA-based middleware tool for distributing and connecting scientific modeling and simulation codes.

Research Experience (Academic)

Doctoral Research — University of New Mexico, 8/01–5/04

Conceived novel mathematical analyses to the direct volume rendering light transport equation that enable real time display of highly accurate, linearly interpolated unstructured volumes.

Selected Publications

- Kwan-Liu Ma, Chaoli Wang, Hongfeng Yu, Kenneth Moreland, Jian Huang, and Rob Ross. "Next-Generation Visualization Technologies: Enabling Discoveries at Extreme Scale." *SciDAC Review*, Issue 12, pg. 12–21, Spring 2009.
- John Biddiscombe, Berk Geveci, Ken Martin, Kenneth Moreland, and David Thompson. "Time Dependent Processing in a Parallel Pipeline Architecture." *IEEE Transactions on Visualization and Computer Graphics*, Volume 13, Number 6, pg. 1376–1383, November/December, 2007. DOI=10.1109/TVCG.2007.70600.

- Kenneth Moreland, Lisa Avila, and Lee Ann Fisk. “Parallel Unstructured Volume Rendering in ParaView.” In *Visualization and Data Analysis 2007, Proceedings of SPIE-IS&T Electronic Imaging*, pg 64950F-1–12. January 2007.
- Andy Cedilnik, Berk Geveci, Kenneth Moreland, James Ahrens, and Jean Farve. “Remote Large Data Visualization in the ParaView Framework.” In *Eurographics Parallel Graphics and Visualization 2006*, pg. 163–170. May 2006.
- Kenneth Moreland and David Thompson. “From Cluster to Wall with VTK.” In *Proceedings of IEEE 2004 Symposium on Parallel and Large-Data Visualization and Graphics*, pages 25–31. October 2003.
- Kenneth Moreland, Brian Wylie, and Constantine Pavlakos. “Sort-Last Parallel Rendering for Viewing Extremely Large Data Sets on Tile Displays.” In *Proceedings of the IEEE Symposium on Parallel and Large-Data Visualization and Graphics*, pages 85–92. October 2001.
- Brian Wylie, Constantine Pavlakos, Vasily Lewis, and Kenneth Moreland. “Scalable Rendering on PC Clusters.” *IEEE Computer Graphics and Applications*, volume 21, number 4, pages 62–70. July/August 2001.

Selected Presentations

- Kenneth Moreland, Andrew Bauer, Pat Marion, and Nathan Fabian. "In-Situ Visualization with the ParaView Coprocessing Library." Tutorial *Supercomputing 2010*, November 2010.
- Kenneth Moreland, John Greenfield, W. Alan Scott, Utkarsh Ayachit, and Berk Geveci. "Large Scale Visualization with ParaView." Tutorial *Supercomputing 2009*, November 2009.
- Kenenth Moreland and David DeMarle. "Parallel Distributed-Memory Visualization with ParaView." Tutorial *IEEE Cluster 2009*, August 2009.
- Kenneth Moreland, John Greenfield, W. Alan Scott, Utkarsh Ayachit, Berk Geveci, and David DeMarle. "Large Scale Visualization with ParaView." *Supercomputing 2008*, November 2008.
- Kenneth Moreland and John Greenfield. "Large Scale Visualization with ParaView 3". *Supercomputing 2007*. November 2007.
- “Parallel Visualization with ParaView.” *Supercomputing 2005*. November 2005.
- “Big Data, Big Displays, and Cluster-Driven Interactive Visualization.” *Workshop on Commodity-Based Visualization Clusters*. October 2002.

Professional Affiliations

- Institute of Electrical and Electronic Engineers (IEEE)
- Association for Computing Machinery (ACM)

Résumé: John Greenfield

Education

- Doctor of Philosophy Computer Engineering, University of New Mexico, Albuquerque, NM, December 2003.
- Master of Science Computer Engineering, University of New Mexico, Albuquerque, NM, May 1992.
- Bachelor of Science Electrical Engineering, University of New Mexico, Albuquerque, NM, May 1985.

Work Experience

Visualization Support ASAP under contract to Sandia National Laboratories,

3/2004 – Present

Provide support to both developers and users of visualization support tools especially EnSight and ParaView.

Sr Research Scientist University of New Mexico High Performance Computing Center,

8/1997 – 3/2003

Researched parallel processing systems and visualization software, Supported large scale videoconferencing systems and user training.

Research Experience (Academic)

Doctoral Research University of New Mexico, 8/2001 – 10/2003

Investigated impact of directional audio on large scale videoconferencing systems.

Selected Publications

- Kenneth L. Summers, John Greenfield and Brian T. Smith. “A Survey of Parallel Program Performance Evaluation Techniques using Visualization and Virtual Reality," in Proceedings, IEEE Aerospace conference 2000.
- Robert A. Ballance, Thomas P. Caudell, John Greenfield and Rick Stevens. “The National Computational Science Alliance Access Grid: An Internet-Based Collaboration Tool Augmented by High-Performance Computing," in Proceedings of the DoD High Performance Computing Users Group Conference 2000, June 2, 2000
- S. Salvini, B.T. Smith, and J. Greenfield, “Towards mixed mode parallelism on the new model F50-Based IBM SP system," Albuquerque High Performance

Computing Center, University of New Mexico, Technical Report AHPCC98003, September 1998.
- John A. Greenfield, Richard D. Hunt, and Gregory L. Heileman. "A Parallel Load Balancing Technique for Finite Element Method Problems," University of New Mexico, Technical Report EECE95-002, February 28, 1995

Selected Presentations

- Large Scale Visualization with ParaView 3 presented at Supercomputing 2007.
- Beginning EnSight class at Sandia National Laboratories 05/01/2007
- Beginning ParaView class at Sandia National Laboratories, 02/27/2007
- "What's New in ParaView" presented at DOE Computer Graphics Forum 2005, 04/2005

Professional Affiliations

- Association for Computing Machinery (ACM)

Résumé: Berk Geveci

Education

- Doctor of Philosophy Mechanical Engineering from Lehigh University, 1999.
- Master of Science Mechanical Engineering from Lehigh University, 1996.
- Bachelor of Science Mechanical Engineering from Bogazici University, 1994.

Work Experience

ParaView Project Lead — Kitware Inc.

Lead the open source ParaView project.

Research Experience (Academic)

Post-Doctoral Fellowship — University of Pennsylvania

Worked in the area of optimal control investigating applications in the control of hydrothermal instabilities.

Doctoral Research — Lehigh University

Conducted research on subsonic and supersonic flow induced nonlinear vibrations, developing a new procedure for the solution of coupled flow and structural equations. In addition, he authored software for the study of separation in unsteady boundary layer flows and the visualization of the numerical and experimental results.

Selected Publications

- John Biddiscombe, Berk Geveci, Ken Martin, Kenneth Moreland, and David Thompson. "Time Dependent Processing in a Parallel Pipeline Architecture." *IEEE Transactions on Visualization and Computer Graphics*, Volume 13, Number 6, pg. 1376–1383, November/December, 2007. DOI=10.1109/TVCG.2007.70600.
- Andy Cedilnik, Berk Geveci, Kenneth Moreland, James Ahrens, and Jean Farve. “Remote Large Data Visualization in the ParaView Framework.” In *Eurographics Parallel Graphics and Visualization 2006*, pg. 163–170. May 2006.
- James Ahrens, Berk Geveci and Charles Law. “ParaView: An End-User Tool for Large Data Visualization.” In *The Visualization Handbook*. Edited by C.D. Hansen and C.R. Johnson. Elsevier. 2005.
- K.M. Martin, B. Geveci, J. Ahrens, C. Law. “Large Scale Data Visualization Using Parallel Data Streaming.” *IEEE Computer Graphics & Applications*, July 2001.

Résumé: Utkarsh Ayachit

Education

- Master of Science Computer Science from University of Maryland, Baltimore County, 2004.

Work Experience

Research Staff — Kitware Inc.

Contributed to the design and implementation of ParaView.

Research Experience (Academic)

Master's Research — Lehigh University

Explored the use of level-of-detail in 2D flow visualization techniques namely streamline visualizations, vector plots and texture-based techniques.