Simulation-Time
Ultra Scale Visualization

Kwan-Liu Ma
UC Davis
University of California
Outline

• Simulation-Time Visualization

• Parallel Rendering Techniques

• Workshop closing remarks
Rendering Time-Varying Data

- Hundreds to thousands of time steps
- Hundreds of MB to hundreds of GB per time step
- Terabytes now and petabyte is upcoming

- Temporal browsing using a desktop PC
  - $512^3$ or smaller volume data
- Temporal browsing using a cluster or massively parallel supercomputer
  - Data distribution
  - Capacity and processing power
  - Scalability
Techniques

• Data Reduction
  – Spatial and temporal domain encoding
  – multi-resolution representations
  – feature extraction & tracking

• Streaming

• Parallel pipelined rendering

• Parallel I/O

• Simulation-time visualization
Simulation-Time Visualization

- Process data while it is being generated
- Prepare (and reduce) data for visualization and analysis
- All the data are available!
- Transfer the transformed/packed data or images
- Monitor simulation
- Steer visualization and simulation
- But …
Simulation-Time Visualization

• Supercomputer time is precious
  – The cost of visualization calculations
  – The scalability of the visualization code
• Memory space is limited
  – The storage overhead of visualization calculations
• Domain decomposition and data organization are optimized for the simulation code
• The simulation code must talk to the visualization code
• What to save is not known a priori
  – Exploratory visualization?
  – Accumulated knowledge in routine study
  – Machine learning
Simulation-time visualization is the ultimate solution to the ultra-scale data analysis problem

• All the data are there
• We can reduce data early in the scientific discovery process
Earthquake Ground Motion Simulation and Visualization

Work done with Tiankai Tu, Jacobo Bielak, Omar Ghattas, and Dave O’Hallaron
End-to-End Tightly Coupled

- The main objective is to eliminate scalability bottlenecks
- Run simulation pipelines end-to-end in parallel
- Execute all components on the same processors
- Partitioning favors the solver
- Sustained flops increases as the problem size increases
Simulation-Time Visualization

Elements:
- 0.661M Elements
- 9.92M Elements
- 31.3M Elements
- 114M Elements
- 462M Elements
- 1.22B Elements

Performance Metrics:
- Replicating
- Meshing
- Solving
- Visualizing
Simulation-Time Visualization
Presentations

1. **Technical paper**
   
   *From Mesh Generation to Scientific Visualization: An End-to-End Approach to Parallel Supercomputing*

   2:30pm - 3pm  Scalable Systems Software Session

   Tuesday

2. **HPC Analytics Challenge Finalist**

   Remote Runtime Steering of Integrated Terascale Simulation and Visualization

   1:30pm - 2pm  Analytics Challenge

   Tuesday
Parallel Rendering
Parallel Volume Rendering

- Data Distribution
- Rendering
- Compositing in visibility order
- Image Collection

Communication

Volume-space

Image-space partitioning

RBGA data
Load Balancing Problem
Parallel Rendering of Large Unstructured Grid Data

1. Domain decomposition according to a static load balancing scheme used by the renderer
   - Can load be balanced?
   - PRS 1997

2. Domain decomposition used by the simulation
   - Load balancing?
   - PGV 2006
Scalable Rendering Design I

- Random distribution of elements
- 3D space partition tree guiding the rendering
- Early compositing for each projected region
- Low sorting cost and memory requirement
- Overlapping rendering and communication
- 75-90% efficiency using up to 512 processors
- Rendering 18 million cells
Test Results

Cray/SGI T3E

Million of tetrahedra Per second

processors

0.56 million cells
18.2 million cells
Scalable Rendering Design II

- The same domain decomposition used by the simulation code
- **Rendering of large elements is deferred**
- Minimal communication cost
- Rendering of large elements are redistributed among processors according to the image space partitioning.
- 80-95% efficiency using up to 400 processors
- rendering up to 27 billion elements

Blast wave over a reinforced concrete
Scalability Study

100M element unstructured grid
1024x1024 pixel image
Summary

- Terascale visualization capabilities have been developed but are not generally accessible
- Petascale computing is coming!
- Parallel visualization must address every stage of the visualization pipeline
- Simulation-time visualization is feasible and attractive
- GPU computing creates both new opportunities and new challenges
- The Institute for Ultra Scale Visualization will lead the effort to create technologies meeting the next generation visualization challenges
Institute for Ultra Scale Visualization

Research
• Time-Varying Data Visualization
• Multivariate Data Visualization
• Irregular/Unstructured Grid Data Visualization
  (J. Huang, K.-L. Ma, N. Max, H.-W. Shen)
• Parallel Rendering Framework and API
• Parallel I/O Support for Visualization
• Multi-GPU Support for Visualization
• ParaView Extensions for Ultra Scale Visualization
  (G. Humphreys, K. Moreland, J. Owens, R. Ross)

Education

http://IUSV.org    ma@cs.ucdavis.edu
Workshop Program

• Opening 08:10 - 08:30am
• State of the Art 08:30 - 10:00am
• Coffee Break 10:00-10:30am
• Case Studies 10:30-12:00pm
• Lunch Break 12:00-01:30pm
• The Future I 01:30-03:00pm
• Coffee Break 03:00-03:30pm
• The Future II 03:30-05:00pm
State of the Art

• VisIt Visualization Tool, Hank Childs, Lawrence Livermore National Laboratory
• ParaView - Parallel Visualization Application, Kenneth Moreland, Sandia National Laboratories
• VAPOR, John Clyne, National Center for Atmospheric Research

Case Studies

• Leadership End-to-End Computing, Scott Klasky, Oak Ridge National Laboratory
• Petascale Visualization on BGL, Michael Papka, Argonne National Laboratory
• Quantitative and Comparative Visualization Applied to Cosmological Simulations, James Ahrens, Los Alamos National Laboratory
The Future

• *Query-Driven Visualization of Large Data Sets*, Wes Bethel, Lawrence Berkeley National Laboratory

• *Large Data Visualization using Shared Distributed Resources*, Jian Huang, University of Tennessee, Knoxville

• *Navigating Large Data Scalar Volume Data*, Han-Wei Shen, Ohio State University

• *Feature Extraction and Tracking Methods*, Deborah Silver, Rutgers, the State University of New Jersey

• *Topology-Based Analysis of Large Scale Data*, Valerio Pascucci, Lawrence Livermore National Laboratory

• *Massively Parallel Visualization*, Kwan-Liu Ma, University of California at Davis
Acknowledgments

• SC06 Workshop Co-Chairs and Committee
  – Mary Thomas
  – Radha Nandkumar
• NSF ITR program
• DOE SciDAC program
Comments/Questions?

Kwan-Liu Ma
ma@cs.ucdavis.edu
http://www.cs.ucdavis.edu/~ma
http://IUSV.org