S4410: Visualization and Analysis of Petascale Molecular Simulations with VMD

John E. Stone
Theoretical and Computational Biophysics Group
Beckman Institute for Advanced Science and Technology
University of Illinois at Urbana-Champaign
http://www.ks.uiuc.edu/Research/gpu/
S4410, GPU Technology Conference
15:30-16:20, Room LL21E, San Jose Convention Center,
San Jose, CA, Tuesday March 25, 2014
VMD – “Visual Molecular Dynamics”

- Visualization and analysis of:
  - molecular dynamics simulations
  - particle systems and whole cells
  - cryoEM densities, volumetric data
  - quantum chemistry calculations
  - sequence information
- User extensible w/ scripting and plugins
- http://www.ks.uiuc.edu/Research/vmd/
Goal: A Computational Microscope
Study the molecular machines in living cells

Ribosome: target for antibiotics

Poliovirus
VMD Interoperability Serves Many Communities

- VMD 1.9.1 user statistics:
  - 74,933 unique registered users from all over the world
- Uniquely interoperable with a broad range of tools: AMBER, CHARMM, CPMD, DL_POLY, GAMESS, GROMACS, HOOMD, LAMMPS, NAMD, and many more …
- Supports key data types, file formats, and databases, e.g. electron microscopy, quantum chemistry, MD trajectories, sequence alignments, super resolution light microscopy
- Incorporates tools for simulation preparation, visualization, and analysis
NAMD and VMD Use GPUs and Petascale Computing to Meet Computational Biology’s Insatiable Demand for Processing Power.
Large-Size and Long-Timescale MD Simulations Drive VMD Development

Extend VMD to enable large state-of-the-art simulations to be performed “routinely”

- Improve display fidelity and performance
- Improve model building tools
- Enable flexible and rapid analysis of multi-terabyte simulation trajectories
- Enable development of force field parameters for drug compounds
- Adapt VMD file formats and internal data structures for new simulation types
First Simulation of a Virus Capsid (2006)

Satellite Tobacco Mosaic Virus (STMV)

First MD simulation of a complete virus capsid
STMV smallest available capsid structure

STMV simulation, visualization, and analysis pushed us toward GPU computing!

MD showed that STMV capsid collapses without its RNA core

1 million atoms
A huge system for 2006

CUDA GPU-Accelerated Trajectory Analysis and Visualization in VMD

<table>
<thead>
<tr>
<th>VMD GPU-Accelerated Feature or GPU Kernel</th>
<th>Exemplary speedup vs. contemporary 4-core CPU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Molecular orbital display</td>
<td>30x</td>
</tr>
<tr>
<td>Radial distribution function</td>
<td>23x</td>
</tr>
<tr>
<td>Molecular surface display</td>
<td>15x</td>
</tr>
<tr>
<td>Electrostatic field calculation</td>
<td>11x</td>
</tr>
<tr>
<td>Ray tracing w/ shadows, AO lighting</td>
<td>7x</td>
</tr>
<tr>
<td>cryoEM cross correlation quality-of-fit</td>
<td>7x</td>
</tr>
<tr>
<td>Ion placement</td>
<td>6x</td>
</tr>
<tr>
<td>MDFF density map synthesis</td>
<td>6x</td>
</tr>
<tr>
<td>Implicit ligand sampling</td>
<td>6x</td>
</tr>
<tr>
<td>Root mean squared fluctuation</td>
<td>6x</td>
</tr>
<tr>
<td>Radius of gyration</td>
<td>5x</td>
</tr>
<tr>
<td>Close contact determination</td>
<td>5x</td>
</tr>
<tr>
<td>Dipole moment calculation</td>
<td>4x</td>
</tr>
</tbody>
</table>
GPU-Accelerated $C_{60}$ Molecular Orbitals

<table>
<thead>
<tr>
<th>Device</th>
<th>CPUs, GPUs</th>
<th>Runtime (s)</th>
<th>Speedup</th>
</tr>
</thead>
<tbody>
<tr>
<td>2x Intel X5550-SSE</td>
<td>8</td>
<td>4.13</td>
<td>1</td>
</tr>
<tr>
<td>GeForce GTX 480</td>
<td>1</td>
<td>0.255</td>
<td>16</td>
</tr>
<tr>
<td>GeForce GTX 480</td>
<td>4</td>
<td>0.081</td>
<td>51</td>
</tr>
</tbody>
</table>

3-D orbital lattice: millions of points

Lattice slices computed on multiple GPUs

GPU threads compute one point

CUDA thread blocks

2-D CUDA grid on each GPU
Multi-GPU RDF Performance

- 4 NVIDIA GTX480 GPUs 30 to 92x faster than 4-core Intel X5550 CPU
- Fermi GPUs ~3x faster than GT200 GPUs: larger on-chip shared memory

Time-Averaged Electrostatics Analysis on Energy-Efficient GPU Cluster

- **1.5 hour** job (CPUs) reduced to **3 min** (CPUs+GPU)
- Electrostatics of thousands of trajectory frames averaged
- Per-node power consumption on NCSA “AC” GPU cluster:
  - CPUs-only: 448 Watt-hours
  - CPUs+GPUs: 43 Watt-hours
- GPU Speedup: **25.5x**
- Power efficiency gain: **10.5x**


2013 *HPCwire* Editors’ Choice Award for Best Use of HPC in Life Sciences
NAMD Titan XK7 Performance August 2013

NAMD XK7 vs. XE6
GPU Speedup: 2x-4x

HIV-1 Trajectory:
~1.2 TB/day
@ 4096 XK7 nodes
VMD Petascale Visualization and Analysis

- Analyze/visualize large trajectories too large to transfer off-site:
  - User-defined parallel analysis operations, data types
  - Parallel rendering, movie making
- Supports GPU-accelerated Cray XK7 nodes for both visualization and analysis:
  - GPU accelerated trajectory analysis w/ CUDA
  - OpenGL and GPU ray tracing for visualization and movie rendering
- Parallel I/O rates up to 275 GB/sec on 8192 Cray XE6 nodes – can read in 231 TB in 15 minutes!

Parallel VMD currently available on:
- ORNL Titan, NCSA Blue Waters, Indiana Big Red II

NCSA Blue Waters Hybrid Cray XE6 / XK7
22,640 XE6 dual-Opteron CPU nodes
4,224 XK7 nodes w/ Telsa K20X GPUs
Structural Route to the all-atom HIV-1 Capsid

1st TEM (1999) 1st tomography (2003) Crystal structures of separated hexamer and pentamer

High res. EM of hexameric tubule, tomography of capsid, all-atom model of capsid by MDFF w/ NAMD & VMD, NSF/NCSA Blue Waters computer at Illinois

Briggs et al. *EMBO J*, 2003
Briggs et al. *Structure*, 2006


cryo-ET (2006)

hexameric tubule

Li et al., *Nature*, 2000
Byeon et al., *Cell* 2009

Molecular Dynamics Flexible Fitting (MDFF)

Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics.

Evaluating Quality-of-Fit for Structures Solved by Hybrid Fitting Methods

Compute Pearson correlation to evaluate the fit of a reference cryo-EM density map with a simulated density map produced from an all-atom structure.
GPUs Can Reduce Trajectory Analysis Runtimes from Hours to Minutes

GPUs enable laptops and desktop workstations to handle tasks that would have previously required a cluster, or a very long wait...

GPU-accelerated petascale supercomputers enable analyses were previously impossible, allowing detailed study of very large structures such as viruses

GPU-accelerated MDFF Cross Correlation Timeline

Regions with poor fit

Regions with good fit
Padding optimizes global memory performance, guaranteeing coalesced global memory accesses.

3-D density map decomposes into 3-D grid of 8x8x8 tiles containing CC partial sums and local CC values.

Small 8x8x2 CUDA thread blocks afford large per-thread register count, shared memory.

Each thread computes 4 z-axis density map lattice points and associated CC partial sums.

Spatial CC map and overall CC value computed in a single pass.

Threads producing results that are used.

Inactive threads, region of discarded output.

Grid of thread blocks.
## VMD GPU Cross Correlation Performance

<table>
<thead>
<tr>
<th>Resolution (Å)</th>
<th>RHDV</th>
<th>Mm-cpn open</th>
<th>GroEL</th>
<th>Aquaporin</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.5</td>
<td>8</td>
<td>4</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Atoms</td>
<td>702K</td>
<td>61K</td>
<td>54K</td>
<td>1.6K</td>
</tr>
<tr>
<td>VMD-CUDA Quadro K6000</td>
<td>0.458s (34.6x)</td>
<td>0.06s (25.7x)</td>
<td>0.034s (36.8x)</td>
<td>0.007s (55.7x)</td>
</tr>
<tr>
<td>VMD-CPU-SSE 32-threads, 2x Xeon E5-2687W</td>
<td>0.779s (20.3x)</td>
<td>0.085s (18.1x)</td>
<td>0.159s (7.9x)</td>
<td>0.033s (11.8x)</td>
</tr>
<tr>
<td>Chimera 1-thread Xeon E5-2687W</td>
<td>15.86s (1.0x)</td>
<td>1.54s (1.0x)</td>
<td>1.25s (1.0x)</td>
<td>0.39s (1.0x)</td>
</tr>
</tbody>
</table>

### VMD RHDV Cross Correlation Timeline on Cray XK7

<table>
<thead>
<tr>
<th></th>
<th>RHDV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Atoms</strong></td>
<td>702K</td>
</tr>
<tr>
<td><strong>Component Selections</strong></td>
<td>720</td>
</tr>
<tr>
<td><strong>Single-node XK7 (projected)</strong></td>
<td>336 hours (14 days)</td>
</tr>
<tr>
<td><strong>128-node XK7</strong></td>
<td>3.2 hours</td>
</tr>
</tbody>
</table>

105x speedup

Calculation would take **5 years** using conventional non-GPU software on a workstation!!
Visualization Goals, Challenges

- Increased GPU acceleration for visualization of **petascale** molecular dynamics trajectories
- **Overcome GPU memory capacity limits**, enable high quality visualization of >100M atom systems
- Use GPU to accelerate not only interactive-rate visualizations, but also photorealistic ray tracing with **artifact-free ambient occlusion lighting**, etc.
- Maintain **ease-of-use**, intimate link to VMD analytical features, atom selection language, etc.
VMD “QuickSurf” Representation, Ray Tracing

All-atom HIV capsid simulations w/ up to 64M atoms on Blue Waters
VMD “QuickSurf” Representation

- Displays continuum of structural detail:
  - All-atom, coarse-grained, cellular models
  - Smoothly variable detail controls
- Linear-time algorithm, scales to millions of particles, as limited by memory capacity
- Uses multi-core CPUs and GPU acceleration to enable smooth interactive animation of molecular dynamics trajectories w/ up to ~1-2 million atoms
- GPU acceleration yields 10x-15x speedup vs. multi-core CPUs

Fast Visualization of Gaussian Density Surfaces for Molecular Dynamics and Particle System Trajectories.
VMD 1.9.2 QuickSurf Algorithm Improvements

- 50%-66% memory use, 1.5x-2x speedup
- Build spatial acceleration data structures, optimize data for GPU
- Compute 3-D density map, 3-D color texture map with **data-parallel “gather” algorithm**:

\[
\rho(\vec{r}, \vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N) = \sum_{i=1}^{N} e^{-\frac{|\vec{r} - \vec{r}_i|^2}{2\alpha^2}}
\]

- Normalize, quantize, and compress density, color, surface normal data **while in registers**, before writing out to GPU global memory
- Extract isosurface, maintaining quantized/compressed data representation

3-D density map lattice, spatial acceleration grid, and extracted surface
VMD GPU-Accelerated Ray Tracing Engine

• Complementary to VMD OpenGL GLSL renderer that uses fast, interactivity-oriented rendering techniques

• **Key ray tracing benefits: ambient occlusion lighting, shadows, high quality transparent surfaces, …**
  
  – Subset of Tachyon parallel ray tracing engine in VMD
  
  – GPU acceleration w/ CUDA+OptiX ameliorates long rendering times associated with advanced lighting and shading algorithms
    
    • Ambient occlusion generates large secondary ray workload
    
    • Transparent surfaces and transmission rays can increase secondary ray counts by another order of magnitude

  – Adaptation of Tachyon to the GPU required careful avoidance of GPU branch divergence, use of GPU memory layouts, etc.
Lighting Comparison, STMV Capsid

Two lights, no shadows

Two lights, hard shadows, 1 shadow ray per light

Ambient occlusion + two lights, 144 AO rays/hit
GPU Ray Tracing of HIV-1 on Blue Waters

- 64M atom simulation, 1079 movie frames
- **Ambient occlusion lighting**, shadows, transparency, antialiasing, depth cueing, **144 rays/pixel minimum**
- GPU memory capacity hurdles:
  - Surface calc. and ray tracing each use **over 75% of K20X 6GB on-board GPU memory** even with quantized/compressed colors, surface normals, ...
  - Evict non-RT GPU data to host prior to ray tracing
  - Eviction was **still required** on a test machine with a **12GB Quadro K6000 GPU** – the multi-pass “QuickSurf” surface algorithm grows the per-pass chunk size to reduce the number of passes
HIV-1 “HD” 1920x1080 movie rendering:
GPUs speed up geom+ray tracing by **up to eight times**

<table>
<thead>
<tr>
<th>Node Type and Count</th>
<th>Script Load Time</th>
<th>State Load Time</th>
<th>Geometry + Ray Tracing</th>
<th>Total Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>256 XE6 CPUs</td>
<td>7 s</td>
<td>160 s</td>
<td>1,374 s</td>
<td>1,541 s</td>
</tr>
<tr>
<td>512 XE6 CPUs</td>
<td>13 s</td>
<td>211 s</td>
<td>808 s</td>
<td>1,032 s</td>
</tr>
<tr>
<td>64 XK7 Tesla K20X GPUs</td>
<td>2 s</td>
<td>38 s</td>
<td>655 s</td>
<td>695 s</td>
</tr>
<tr>
<td>128 XK7 Tesla K20X GPUs</td>
<td>4 s</td>
<td>74 s</td>
<td>331 s</td>
<td>410 s</td>
</tr>
<tr>
<td>256 XK7 Tesla K20X GPUs</td>
<td>7 s</td>
<td>110 s</td>
<td>171 s</td>
<td>288 s</td>
</tr>
</tbody>
</table>

GPU-Accelerated Molecular Visualization on Petascale Supercomputing Platforms,
VMD 1.9.2 Coming Soon

- Handle very large structures:
  - Tested with up to 240M atoms/particles
  - Atom selections: 1.5x to 4x faster
  - QuickSurf surface display: 1.5x to 2x faster
  - GPU ray tracing: 4x-8x faster than CPU

- Improved movie making tools, off-screen OpenGL movie rendering, parallel movie rendering

- Improved structure building tools

- Many new and updated user-contributed plugins:
  - Bendix – intuitive helix visualization and analysis
  - NMWiz – visual analysis of normal modes
  - Topotools – structure preparation, e.g. for LAMMPS

GPU Ray Tracing of HIV-1 Capsid Detail
Force Field Toolkit (ffTK)

**Current Features**
- Optimize charges, bonds, angles, dihedrals
- GUI with a defined modular workflow
- Automation of tedious tasks
- Tools to assess parameter performance

**Planned Features**
- Support multiple QM software packages
- Optimize AMBER parameters

---

*J. Comp. Chem, 34:2757-2770, 2013*
Plans: **Interactive** Ray Tracing of Molecular Graphics

- STMV virus capsid on a **laptop** GeForce GTX 560M
- **Ambient occlusion lighting**, shadows, reflections, transparency, and much more…

![Standard OpenGL rasterization](image1)

![VMD w/ new GPU ray tracing engine based on CUDA + OptiX: 5-10 FPS](image2)
Plans: Extend Analysis Features of VMD

Exemplary features:
• New secondary structure determination algorithm
  o Support large biomolecular complexes
  o Compute and display time-varying secondary structure interactively
• Simplify analysis of multi-terabyte MD trajectories
  o Circumvent storing large trajectories in memory
  o Out-of-core SSD trajectory access: 7.5 GB/sec
• Automate parallelization of user-defined analysis calculations, interfaced to Timeline plugin


Plans: Analyze Long Simulations with Timeline

TimeLine 2D plot

Rho hexameric helicase 3D structure

Timeline:
- graphing and analysis tool to **identify events** in an MD trajectory
- live 2D whole-trajectory plot linked to 3D structure
- user-extendable

- Perform analysis faster
  - High-performance parallel trajectory analysis on supercomputers and clusters
  - Prototypes show **3500x speedup** on Blue Waters
- Analysis types: filtering, time series analysis, sorting (e.g. bond energies)
- Remote interactive analysis: data at supercomputer center; view in office
Plans: Tablet VMD, Improved Touch Interfaces

• Developed first multi-touch VMD interface
  o Early technology development in advance of devices
  o Collaborative multi-user wireless control of VMD session

• Features in development:
  o tablet display of trajectory timelines, sequence data, plots, and tabular information

• Goal: full tablet-native VMD
Optimizing VMD for Power Consumption
Acknowledgements

• Theoretical and Computational Biophysics Group, University of Illinois at Urbana-Champaign
• NVIDIA CUDA Center of Excellence, University of Illinois at Urbana-Champaign
• NVIDIA CUDA team
• NVIDIA OptiX team
• NCSA Blue Waters Team
• Funding:
  – DOE INCITE, ORNL Titan: DE-AC05-00OR22725
  – NSF Blue Waters:
    NSF OCI 07-25070, PRAC “The Computational Microscope”
  – NIH support: 9P41GM104601, 5R01GM098243-02
NIH BTRC for Macromolecular Modeling and Bioinformatics

Beckman Institute
University of Illinois at Urbana-Champaign

1990-2017
• **GPU-Accelerated Molecular Visualization on Petascale Supercomputing Platforms.**  

• **Early Experiences Scaling VMD Molecular Visualization and Analysis Jobs on Blue Waters.**  

• **Lattice Microbes: High-performance stochastic simulation method for the reaction-diffusion master equation.**  

• **Fast Visualization of Gaussian Density Surfaces for Molecular Dynamics and Particle System Trajectories.**  

• **Immersive Out-of-Core Visualization of Large-Size and Long-Timescale Molecular Dynamics Trajectories.**  

• **Fast Analysis of Molecular Dynamics Trajectories with Graphics Processing Units – Radial Distribution Functions.**  
GPU Computing Publications
http://www.ks.uiuc.edu/Research/gpu/


GPU Computing Publications
http://www.ks.uiuc.edu/Research/gpu/


GPU Computing Publications
http://www.ks.uiuc.edu/Research/gpu/


