GPU-Accelerated Analysis of Large Biomolecular Complexes

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/

Supercomputing 2014 Exhibition
New Orleans, LA, November 18, 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/

- Whole Cell Simulation
- MD Simulations
- Sequence Data
- Quantum Chemistry
- CryoEM, Cellular Tomography
Goal: A Computational Microscope
Study the molecular machines in living cells
Ribosome: target for antibiotics
Poliovirus
NAMD and VMD Use GPUs & Petascale Computing to Meet Computational Biology’s Insatiable Demand for Processing Power
Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics.

Molecular Dynamics Flexible Fitting - Theory

Two terms are added to the MD potential

$$U_{total} = U_{MD} + U_{EM} + U_{SS}$$

An external potential derived from the EM map is defined on a grid as

$$U_{EM}(R) = \sum_j w_j V_{EM}(r_j)$$

$$V_{EM}(r) = \begin{cases} 
\xi \left(1 - \frac{\Phi(r) - \Phi_{thr}}{\Phi_{max} - \Phi_{thr}}\right) & \text{if } \Phi(r) \geq \Phi_{thr}, \\
\xi & \text{if } \Phi(r) < \Phi_{thr}.
\end{cases}$$

A mass-weighted force is then applied to each atom

$$f_{i}^{EM} = -\nabla U_{EM}(R) = -w_i \partial V_{EM}(r_i)/\partial r_i$$
Structural Route to the HIV-1 Capsid

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

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

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


hexameric tubules

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

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 MDFF 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 that were previously impractical, allowing detailed study of very large structures such as viruses.

GPU-accelerated MDFF Cross Correlation Timeline

Regions with poor fit

Regions with good fit
3-D density map decomposes into 3-D grid of $8 \times 8 \times 8$ tiles containing CC partial sums and local CC values.

Small $8 \times 8 \times 2$ 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.

Padding optimizes global memory performance, guaranteeing coalesced global memory accesses.

Fusion of density and CC calculations into a single CUDA kernel!!!

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></th>
<th>RHDV</th>
<th>Mm-cpn open</th>
<th>GroEL</th>
<th>Aquaporin</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Resolution (Å)</strong></td>
<td>6.5</td>
<td>8</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td><strong>Atoms</strong></td>
<td>702K</td>
<td>61K</td>
<td>54K</td>
<td>1.6K</td>
</tr>
<tr>
<td><strong>VMD-CUDA Quadro K6000</strong></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><strong>VMD-CPU-SSE 32-threads, 2x Xeon E5-2687W</strong></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><strong>Chimera 1-thread Xeon E5-2687W</strong></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>Atoms</td>
<td>702K</td>
</tr>
<tr>
<td>Traj. Frames</td>
<td>10,000</td>
</tr>
<tr>
<td>Component Selections</td>
<td>720</td>
</tr>
<tr>
<td>Single-node XK7</td>
<td>336 hours (14 days)</td>
</tr>
<tr>
<td>(projected)</td>
<td></td>
</tr>
<tr>
<td>128-node XK7</td>
<td>3.2 hours</td>
</tr>
<tr>
<td></td>
<td>105x speedup</td>
</tr>
<tr>
<td>2048-node XK7</td>
<td>19.5 minutes</td>
</tr>
<tr>
<td></td>
<td>1035x speedup</td>
</tr>
</tbody>
</table>

Calculation would take **5 years** using original serial CC calculation on a workstation!
VMD GPU-Accelerated Ray Tracing

All-atom HIV capsid simulations w/ up to 64M atoms on Blue Waters
Lighting Comparison

- Two lights, no shadows
- Two lights, hard shadows, 1 shadow ray per light
- Ambient occlusion + two lights, 144 AO rays/hit
HIV-1 Parallel HD Movie Rendering on Blue Waters Cray XE6/XK7

New “TachyonL-OptiX” on XK7 vs. Tachyon on XE6: K20X GPUs yield up to eight times geom+ray tracing speedup

<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>

VMD 1.9.2 Interactive GPU Ray Tracing

• Ray tracing heavily used for VMD publication-quality images/movies
• High quality lighting, shadows, transparency, depth-of-field focal blur, etc.
• VMD now provides –interactive– ray tracing on laptops, desktops, and remote visual supercomputers
VMD TachyonL-OptiX: Multi-GPU on a Desktop or Single Node

VMD Scene

TrBvh RT Acceleration Structure

Scene Data Replicated, Image Space Parallel Decomposition onto GPUs

GPU 0

GPU 1

GPU 2

GPU 3
VMD TachyonL-OptiX: Multi-GPU on NVIDIA VCA Cluster

Scene Data Replicated, Image Space / Sample Space Parallel Decomposition onto GPUs

VCA 0:
8 K6000 GPUs

VCA N:
8 K6000 GPUs
VMD TachyonL-OptiX: Multi-GPU on NVIDIA VCA Cluster

See the live demos!
Acknowledgements

• Theoretical and Computational Biophysics Group, University of Illinois at Urbana-Champaign

• NVIDIA CUDA Center of Excellence, University of Illinois at Urbana-Champaign

• NVIDIA OptiX Team

• Funding:
  – NIH support: 9P41GM104601, 5R01GM098243-02
  – NSF PRAC “The Computational Microscope”, OCI-0832673 and ACI-1440026, and Blue Waters OCI 07-25070 and ACI-1238993
  – DOE INCITE DE-AC05-00OR22725
NIH BTRC for Macromolecular Modeling and Bioinformatics

Beckman Institute
University of Illinois at Urbana-Champaign

1990-2017
GPU Computing Publications
http://www.ks.uiuc.edu/Research/gpu/

SC'14 Visualization and Data Analytics Showcase, 2014. (In press)
Winner of the SC'14 Visualization and Data Analytics Showcase


• Unlocking the Full Potential of the Cray XK7 Accelerator. M. D. Klein and J. E. Stone.
Cray Users Group, Lugano Switzerland, 2014. (In press)


• Simulation of reaction diffusion processes over biologically relevant size and time scales using multi-GPU workstations. M. J. Hallock, J. E. Stone, E. Roberts, C. Fry, and Z. Luthey-Schulten.