# 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,

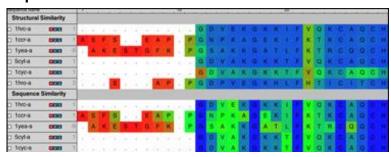
San Jose, CA, Tuesday March 25, 2014





#### VMD - "Visual Molecular Dynamics"

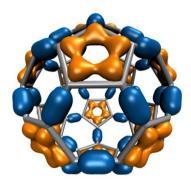
- Visualization and analysis of:
  - molecular dynamics simulations
  - particle systems and whole cells
  - cryoEM densities, volumetric data
  - quantum chemistry calculations
  - sequence information
- User extensible w/ scripting and plugins
- http://www.ks.uiuc.edu/Research/vmd/



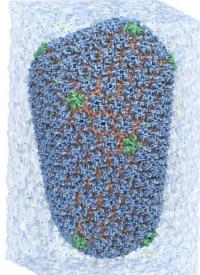
Sequence Data



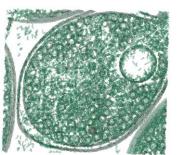
Whole Cell Simulation



**Quantum Chemistry** 



**MD Simulations** 

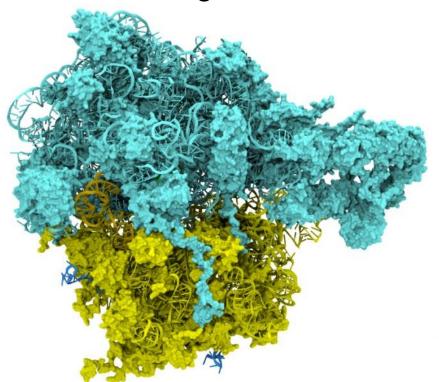


CryoEM, Cellular Tomography

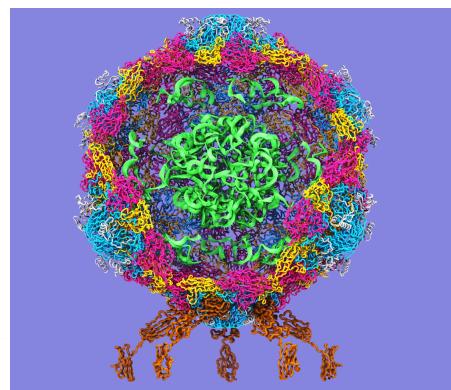
#### Goal: A Computational Microscope

Study the molecular machines in living cells

Ribosome: target for antibiotics

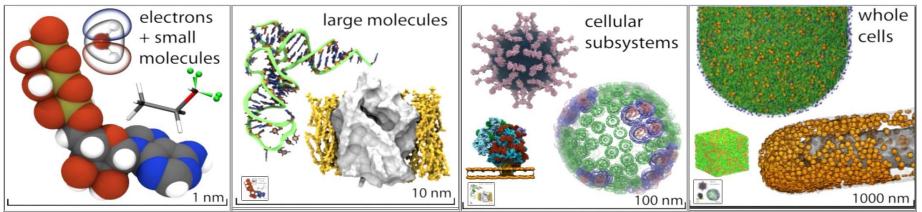


**Poliovirus** 

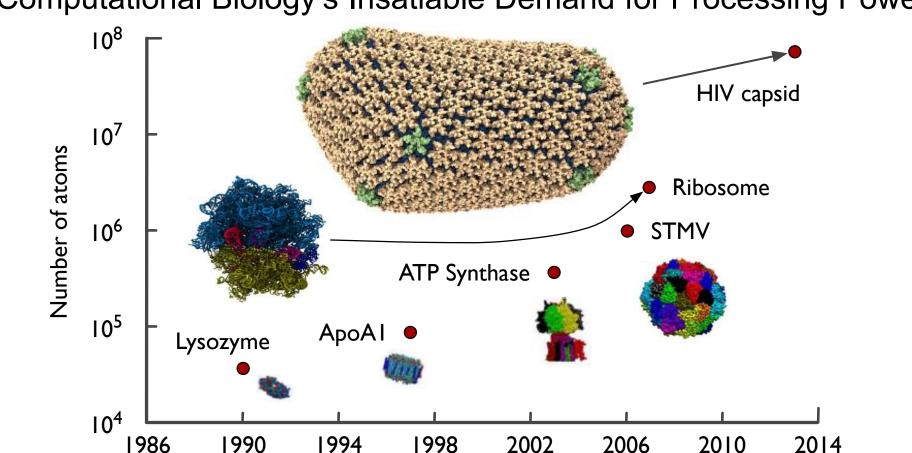


#### VMD Interoperability Serves Many Communities

- 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



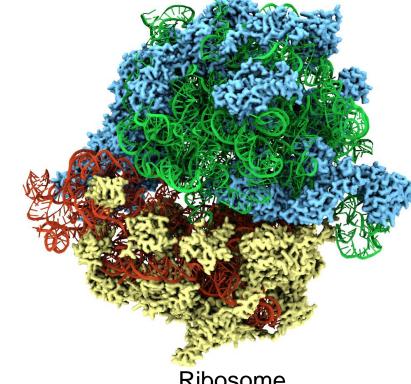
## NAMD and VMD Use GPUs and Petascale Computing to Meet Computational Biology's Insatiable Demand for Processing Power



#### 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 multiterabyte simulation trajectories
- Enable development of force field parameters for drug compounds
- Adapt VMD file formats and internal data structures for new simulation types



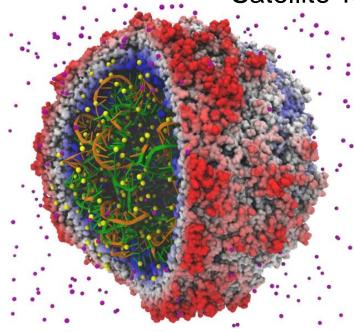






### First Simulation of a Virus Capsid (2006)

Satellite Tobacco Mosaic Virus (STMV)

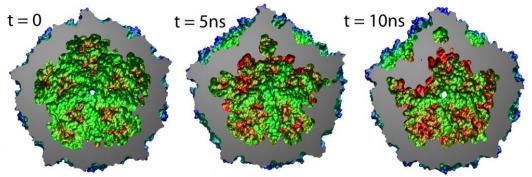


1 million atoms
A huge system for 2006

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

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

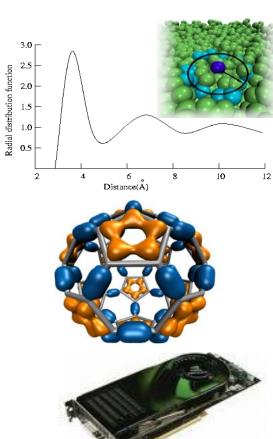
MD showed that STMV capsid collapses without its RNA core



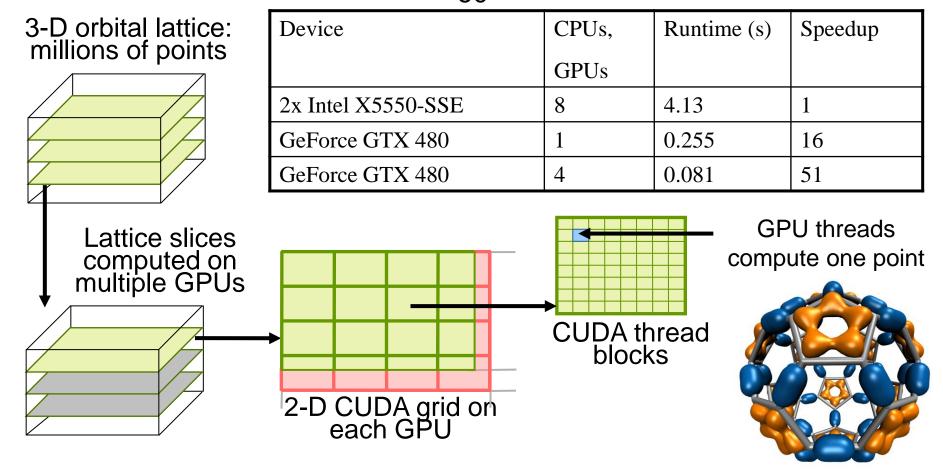
Freddolino et al., Structure, 14:437 (2006)

## CUDA GPU-Accelerated Trajectory Analysis and Visualization in VMD

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

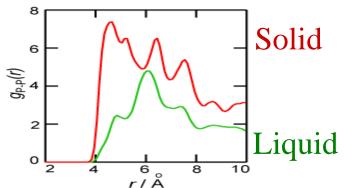


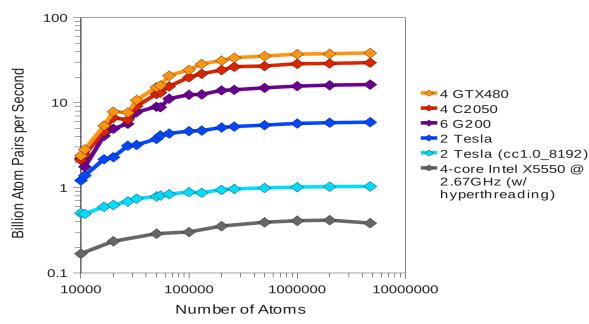
#### GPU-Accelerated C<sub>60</sub> Molecular Orbitals



#### 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





Fast Analysis of Molecular Dynamics Trajectories with Graphics Processing Units – Radial Distribution Functions.

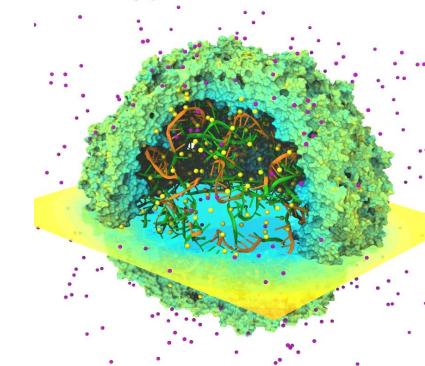
Levine, et al., *J. Comp. Physics*, 230(9):3556-3569, 2011.



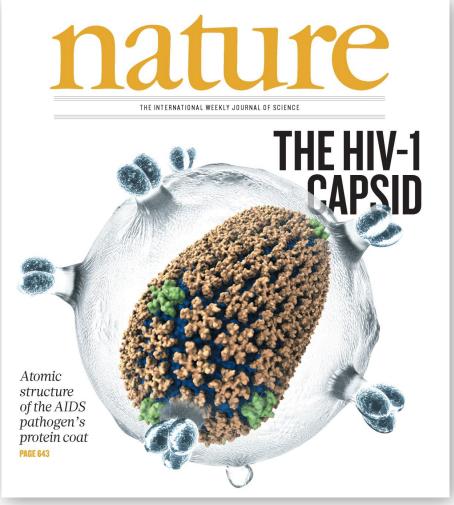


#### Time-Averaged Electrostatics Analysis on Energy-Efficient GPU Cluşter

- 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: **25.5**x
- Power efficiency gain: 10.5x



Quantifying the Impact of GPUs on Performance and Energy Efficiency in HPC Clusters. Enos, et al. *The Work in Progress in Green Computing*, pp. 317-324, 2010.



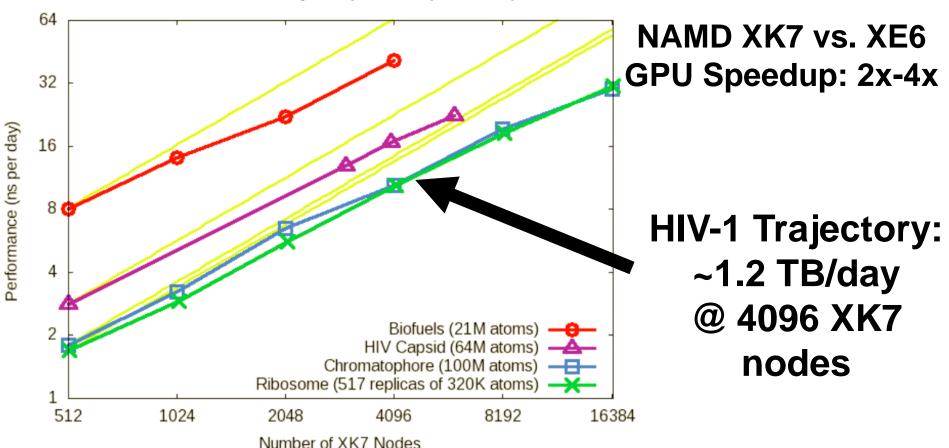
2013 HPCwire
Editors' Choice Award
for Best Use of HPC
in Life Sciences



Zhao et al., Nature 497: 643-646 (2013)

#### NAMD Titan XK7 Performance August 2013

NAMD on Titan Cray XK7 (2fs timestep with PME)



#### 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



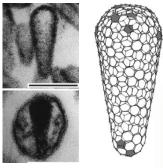
NCSA Blue Waters Hybrid Cray XE6 / XK7 22,640 XE6 dual-Opteron CPU nodes 4,224 XK7 nodes w/ Telsa K20X GPUs



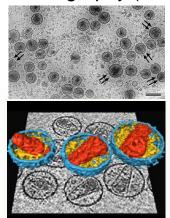


#### Structural Route to the all-atom HIV-1 Capsid

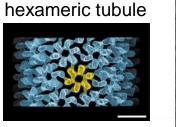
1st TEM (1999) 1st tomography (2003)



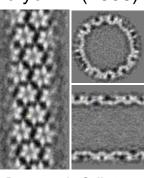
Ganser et al. *Science*, 1999 Briggs et al. *EMBO J*, 2003 Briggs et al. *Structure*, 2006



cryo-ET (2006)

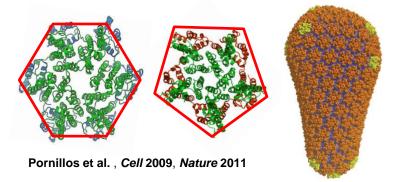


Li et al., Nature, 2000

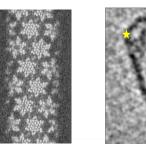


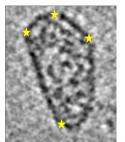
Byeon et al., Cell 2009

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

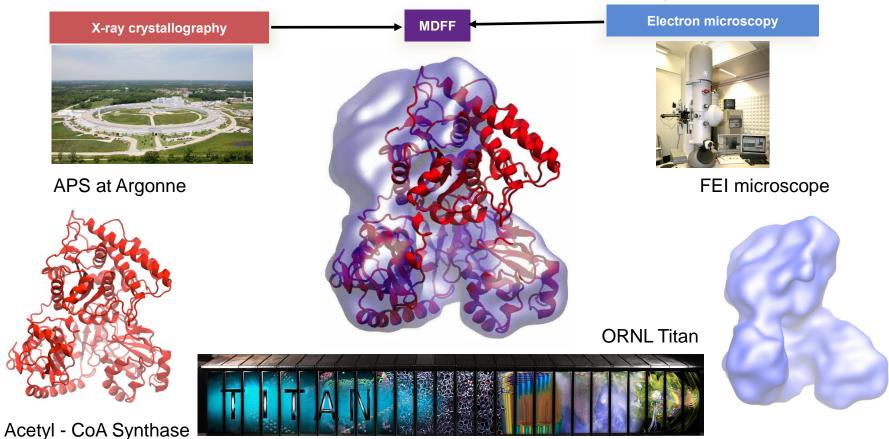






Zhao et al., *Nature* 497: 643-646 (2013)

#### Molecular Dynamics Flexible Fitting (MDFF)

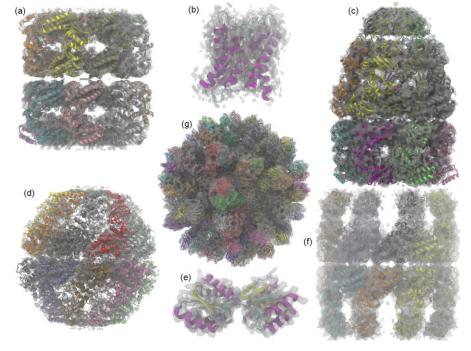






# Evaluating Quality-of-Fit for Structures Solved by Hybrid Fitting Methods

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.



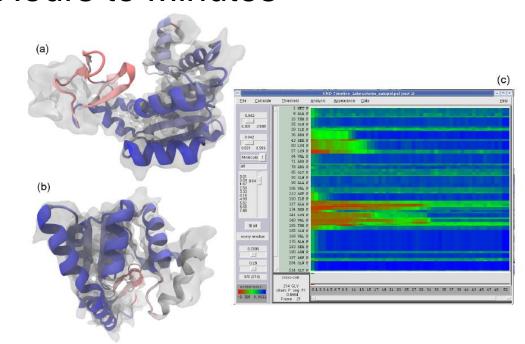




## GPUs Can Reduce Trajectory Analysis Runtimes from Hours to Minutes

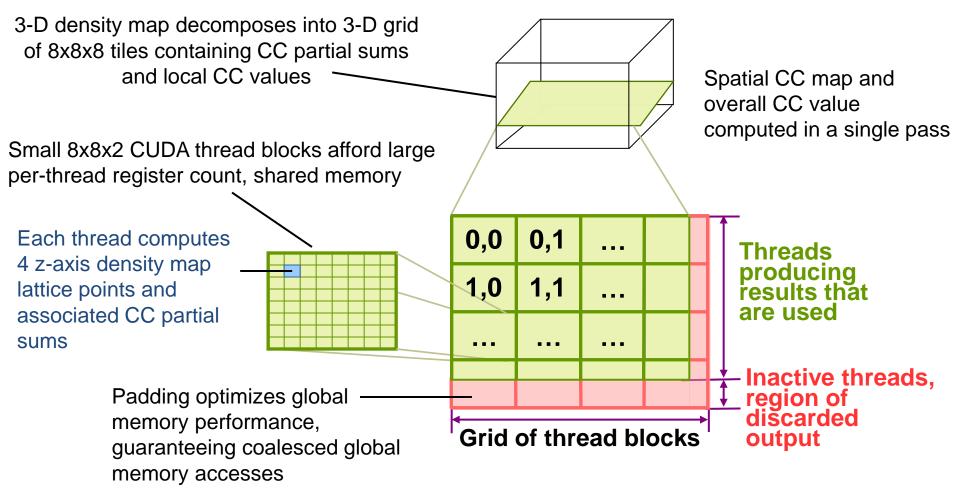
GPUs enable laptops and desktop workstations to handle tasks that would have previously required a cluster, or a very long wait...

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

#### Single-Pass MDFF GPU Cross-Correlation



#### VMD GPU Cross Correlation Performance

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

GPU-accelerated analysis and visualization of large structures solved by molecular dynamics flexible fitting. J. E. Stone, R. McGreevy, B. Isralewitz, and K. Schulten. Faraday Discussion 169, 2014. (In press).



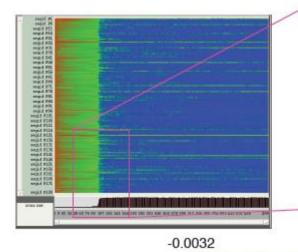


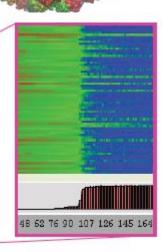
# VMD RHDV Cross Correlation Timeline on Cray XK7

	RHDV
Atoms	702K
Component Selections	720
Single-node XK7 (projected)	336 hours (14 days)
128-node XK7	3.2 hours 105x speedup

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







0.02





#### 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



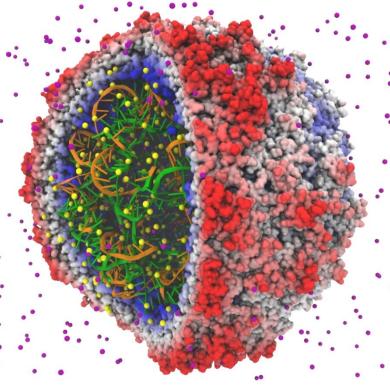
All-atom HIV capsid simulations w/ up to 64M atoms on Blue Waters

#### 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.

M. Krone, J. E. Stone, T. Ertl, K. Schulten. *EuroVis Short Papers*, pp. 67-71, 2012



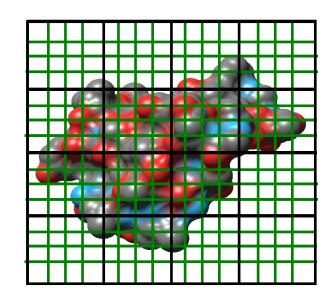
**Satellite Tobacco Mosaic Virus** 

#### 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:

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

- 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.

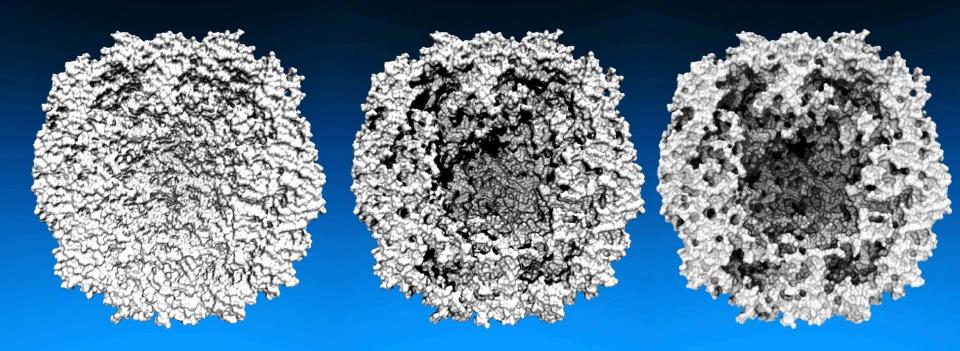




#### Lighting Comparison, STMV Capsid

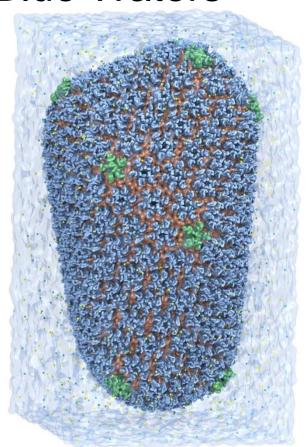
Two lights, no shadows

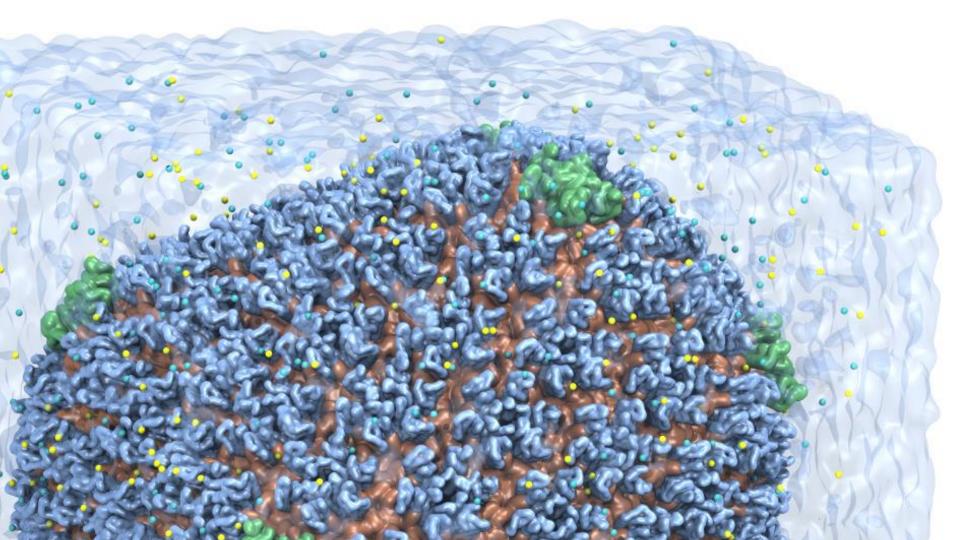
Two lights, hard shadows, <u>1 shadow ray per light</u> Ambient occlusion + two lights, 144 AO rays/hit



#### GPU Ray Tracing of HIV-1 on Blue Waters

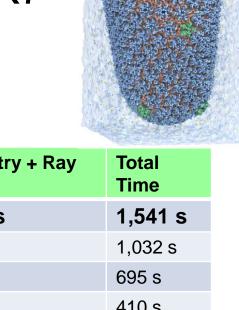
- 64M atom simulation, 1079 movie frames
- Ambient occlusion lighting, shadows, transparency, antialiasing, depth cueing,
   144 rays/pixel minimum
- GPU memory capacity hurdles:
  - Surface calc. and ray tracing each use over 75% of K20X 6GB on-board GPU memory even with quantized/compressed colors, 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





# VMD HIV-1 Parallel Movie Rendering 1.9.2 on Blue Waters Cray XE6/XK7

HIV-1 "HD" 1920x1080 movie rendering: GPUs speed up geom+ray tracing by **up to eight times** 

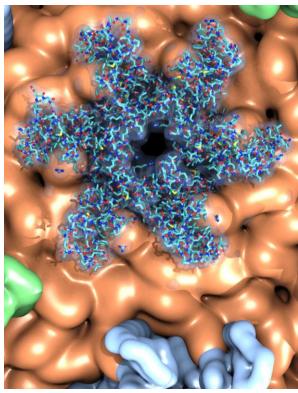


Node Type and Count	Script Load Time	State Load Time	Geometry + Ray Tracing	Total Time
256 XE6 CPUs	7 s	160 s	1,374 s	1,541 s
512 XE6 CPUs	13 s	211 s	808 s	1,032 s
64 XK7 Tesla K20X GPUs	2 s	38 s	655 s	695 s
128 XK7 Tesla K20X GPUs	4 s	74 s	331 s	410 s
256 XK7 Tesla K20X GPUs	7 s	110 s	171 s	288 s

**GPU-Accelerated Molecular Visualization on Petascale Supercomputing Platforms,** Stone et al. UltraVis'13: Eighth Workshop on Ultrascale Visualization Proceedings, 2013.

### 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
  - o 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

#### **VMD**

### Force Field Toolkit (ffTK)

1.9.2

rapid parameterization of small molecules

#### **Current Features**

Optimize charges, bonds, angles, dihedrals

GUI with a defined modular workflow

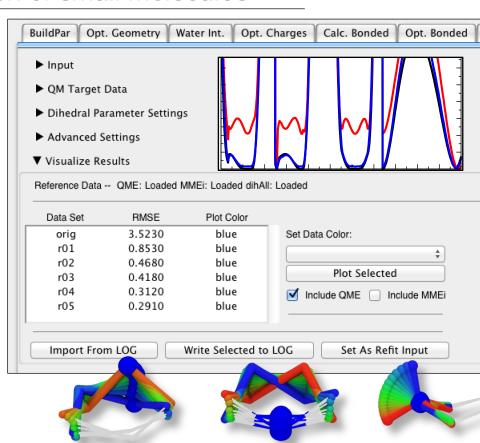
Automation of tedious tasks

Tools to assess parameter performance

#### **Planned Features**

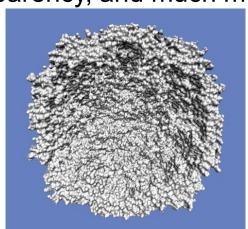
Support multiple QM software packages
Optimize AMBER parameters

J. Comp. Chem, 34:2757-2770, 2013

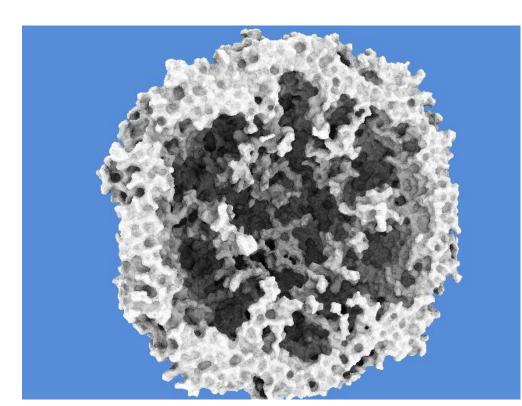


#### Plans: Interactive Ray Tracing of Molecular Graphics

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



Standard OpenGL rasterization



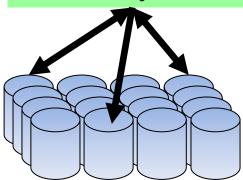
VMD w/ new GPU ray tracing engine based on CUDA + OptiX: 5-10 FPS

#### Plans: Extend Analysis Features of VMD

#### Exemplary features:

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

## Parallel VMD Analysis



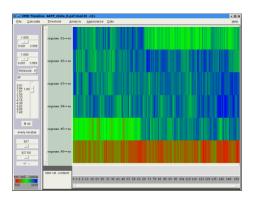
Tens to hundreds of multi-TB trajectories

Immersive Out-of-Core Visualization of Large-Size and Long-Timescale Molecular Dynamics

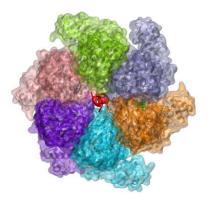
**Trajectories.** J. Stone, K. L. Vandivort, and K. Schulten. G. Bebis et al. (Eds.): 7th International Symposium on Visual Computing (ISVC 2011), LNCS 6939, pp. 1-12, 2011.

Early Experiences Scaling VMD Molecular Visualization and Analysis Jobs on Blue Waters, Proceedings of the Extreme Scaling Workshop, Stone, et al., XSEDE Extreme Scaling Workshop, 2013.

#### Plans: Analyze Long Simulations with Timeline



TimeLine 2D plot



Rho hexameric helicase 3D structure

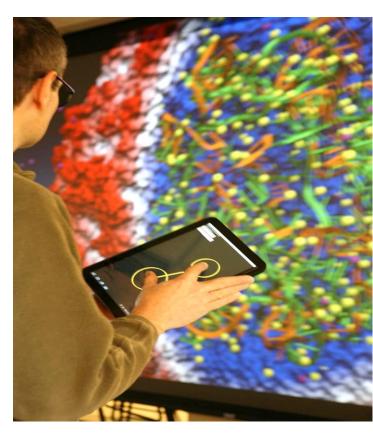
#### Timeline:

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

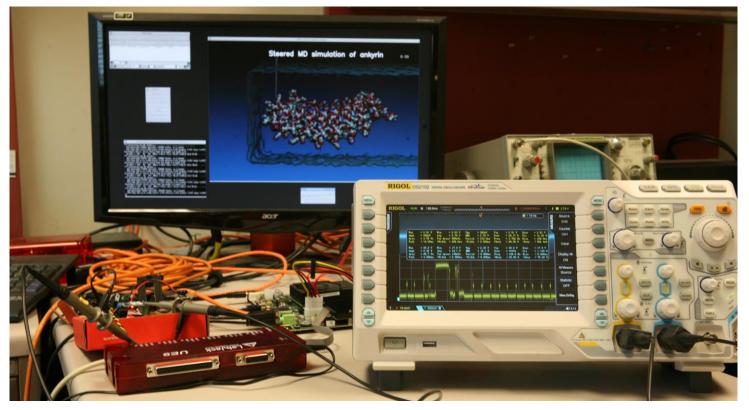
- Perform analysis faster
  - High-performance parallel trajectory analysis on supercomputers and clusters
  - Prototypes show 3500x speedup on Blue Waters
- Analysis types: filtering, time series analysis, sorting (e.g. bond energies)
- Remote interactive analysis: data at supercomputer center; view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

#### Plans: Tablet VMD, Improved Touch Interfaces

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











- **GPU-Accelerated Molecular Visualization on Petascale Supercomputing Platforms.**J. Stone, K. L. Vandivort, and K. Schulten. UltraVis'13: Proceedings of the 8th International Workshop on Ultrascale Visualization, pp. 6:1-6:8, 2013.
- Early Experiences Scaling VMD Molecular Visualization and Analysis Jobs on Blue Waters.

  J. Stone, B. Isralewitz, and K. Schulten. In proceedings, Extreme Scaling Workshop, 2013.
- Lattice Microbes: High-performance stochastic simulation method for the reaction-diffusion master equation. E. Roberts, J. Stone, and Z. Luthey-Schulten.
   J. Computational Chemistry 34 (3), 245-255, 2013.
- Fast Visualization of Gaussian Density Surfaces for Molecular Dynamics and Particle System Trajectories. M. Krone, J. Stone, T. Ertl, and K. Schulten. *EuroVis Short Papers*, pp. 67-71, 2012.
- Immersive Out-of-Core Visualization of Large-Size and Long-Timescale Molecular Dynamics Trajectories. J. Stone, K. L. Vandivort, and K. Schulten. G. Bebis et al. (Eds.): 7th International Symposium on Visual Computing (ISVC 2011), LNCS 6939, pp. 1-12, 2011.
- Fast Analysis of Molecular Dynamics Trajectories with Graphics Processing Units Radial Distribution Functions. B. Levine, J. Stone, and A. Kohlmeyer. *J. Comp. Physics*, 230(9):3556-3569, 2011.





- Quantifying the Impact of GPUs on Performance and Energy Efficiency in HPC Clusters.
  - J. Enos, C. Steffen, J. Fullop, M. Showerman, G. Shi, K. Esler, V. Kindratenko, J. Stone, J. Phillips. *International Conference on Green Computing*, pp. 317-324, 2010.
- **GPU-accelerated molecular modeling coming of age.** J. Stone, D. Hardy, I. Ufimtsev, K. Schulten. *J. Molecular Graphics and Modeling*, 29:116-125, 2010.
- OpenCL: A Parallel Programming Standard for Heterogeneous Computing.
   J. Stone, D. Gohara, G. Shi. Computing in Science and Engineering, 12(3):66-73, 2010.
- An Asymmetric Distributed Shared Memory Model for Heterogeneous Computing Systems. I. Gelado, J. Stone, J. Cabezas, S. Patel, N. Navarro, W. Hwu. *ASPLOS '10: Proceedings of the 15<sup>th</sup> International Conference on Architectural Support for Programming Languages and Operating Systems*, pp. 347-358, 2010.





- **GPU Clusters for High Performance Computing**. V. Kindratenko, J. Enos, G. Shi, M. Showerman, G. Arnold, J. Stone, J. Phillips, W. Hwu. *Workshop on Parallel Programming on Accelerator Clusters (PPAC)*, In Proceedings IEEE Cluster 2009, pp. 1-8, Aug. 2009.
- Long time-scale simulations of in vivo diffusion using GPU hardware. E. Roberts, J. Stone, L. Sepulveda, W. Hwu, Z. Luthey-Schulten. In IPDPS'09: Proceedings of the 2009 IEEE International Symposium on Parallel & Distributed Computing, pp. 1-8, 2009.
- High Performance Computation and Interactive Display of Molecular Orbitals on GPUs and Multi-core CPUs. J. Stone, J. Saam, D. Hardy, K. Vandivort, W. Hwu, K. Schulten, 2nd Workshop on General-Purpose Computation on Graphics Pricessing Units (GPGPU-2), ACM International Conference Proceeding Series, volume 383, pp. 9-18, 2009.
- **Probing Biomolecular Machines with Graphics Processors**. J. Phillips, J. Stone. *Communications of the ACM*, 52(10):34-41, 2009.
- Multilevel summation of electrostatic potentials using graphics processing units. D. Hardy, J. Stone, K. Schulten. *J. Parallel Computing*, 35:164-177, 2009.





- Adapting a message-driven parallel application to GPU-accelerated clusters.
   J. Phillips, J. Stone, K. Schulten. *Proceedings of the 2008 ACM/IEEE Conference on Supercomputing*, IEEE Press, 2008.
- **GPU acceleration of cutoff pair potentials for molecular modeling applications**. C. Rodrigues, D. Hardy, J. Stone, K. Schulten, and W. Hwu. *Proceedings of the 2008 Conference On Computing Frontiers*, pp. 273-282, 2008.
- **GPU computing**. J. Owens, M. Houston, D. Luebke, S. Green, J. Stone, J. Phillips. *Proceedings of the IEEE*, 96:879-899, 2008.
- Accelerating molecular modeling applications with graphics processors. J. Stone, J. Phillips, P. Freddolino, D. Hardy, L. Trabuco, K. Schulten. *J. Comp. Chem.*, 28:2618-2640, 2007.
- Continuous fluorescence microphotolysis and correlation spectroscopy. A. Arkhipov, J. Hüve, M. Kahms, R. Peters, K. Schulten. *Biophysical Journal*, 93:4006-4017, 2007.



