ATPESC Track 7: Scalable Molecular Visualization and Analysis Tools in 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/vmd/ Argonne Training Program on Exascale Computing (ATPESC) 3:30pm-4:30pm, St. Charles Amphitheater, Q Center, St. Charles, IL, Thursday August 9th, 2018



Biomedical Technology Research Center for Macromolecular Modeling and Bioinformatics Beckman Institute, University of Illinois at Urbana-Champaign - www.ks.uiuc.edu



VMD – "Visual Molecular Dynamics"

- 100,000 active users worldwide
- Visualization and analysis of:
 - Molecular dynamics simulations
 - Lattice cell simulations
 - Quantum chemistry calculations
 - Cryo-EM densities, volumetric data
- User extensible scripting and plugins
- http://www.ks.uiuc.edu/Research/vmd/



Cell-Scale Modeling





VMD Hands-On Tutorials

- http://www.ks.uiuc.edu/Training/Tutorials/#vmd
 - Main VMD tutorial
 - QwikMD simulation preparation and analysis plugin
 - VMD images and movies tutorial
 - Structure check
 - VMD quantum chemistry visualization tutorial
 - Visualization and analysis of CPMD data with VMD
 - Parameterizing small molecules using ffTK





VMD Interoperability Serves Many Communities

- 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
- Incorporates tools for simulation preparation, visualization, and analysis



VMD Interoperates with Mainstream Research Tools

- Provides tools for simulation preparation, visualization, and analysis
- Interpret and process multi-modal structural information
- Connects with key software tools to enable state-of-the-art simulations
- Openness, extensibility, and interoperability are VMD hallmarks
- Uses advanced algorithms and hardware technologies to address data size challenges posed by cutting-edge experimental imaging and simulation



Goal: A Computational Microscope

Study the molecular machines in living cells

Ribosome: target for antibiotics

Poliovirus





VMD Approach to Visualization

- Molecular scene is composed of "graphical representations"
- Each representation encapsulates a group of selected atoms, a drawing style, coloring style, and other parameters
- Representations are independent of each other, can be toggled on/off easily, allowing molecular scenes to be builtup incrementally
- VMD atom selection language is shared with its analytical and scripting interfaces

Selection, Filtering

- Most viz tools allow interactive visual picking, menu-driven selections of structure components to display or operate on
- VMD also extensively uses a text-based selection language (think google):

water within 10 of protein and z > 0 nucleic or protein or ions segname BR name "C.*"

- Allows selection on user-defined data fields
- Promotes synergy between interactive and scripting interfaces, visualization and quantitative analysis tasks
- Works well with huge time-varying structures



Structure Visualization

Molecular representations provide different levels of abstraction, atomic detail vs. higher level organizational information

- Atoms, VdW spheres, bonds, ball-stick, ...
- Coarse-grained "beads"
- Ribbons, secondary structure, "cartoon" reps, RNA/DNA
- Molecular surfaces
- Molecular orbitals (quantum chemistry)



Computed Properties

- Smoothing of thermal noise
- Secondary structure
- Hydrogen bonds, salt bridges
- Forces, energies, stress, strain
- Time averaging of electrostatic fields, occupancy maps
- Quality-of-fit cross correlation with cryo-EM density maps
- Normal modes, principal component analysis, essential dynamics
- Cluster simulation trajectory timesteps by structural similarity



Chemoreceptor trimer-ofdimers analysis with Bendix plugin in VMD

Display of Computed Properties on Structures



Per-residue solvent-accessible surface area of Ubiquitin

PME electrostatic potential contour for a helicase on a volumetric slice plane

CheA kinase PCA: first principal component porcupine plot









Computing Molecular Properties

Compute properties, e.g., density, distance, occupancy, electrostatic potential over thousands of MD simulation trajectory frames

Example: display binding sites for diffusively bound ions as probability density isosurfaces

Visualization of Molecular Dynamics

- Molecular dynamics simulations save trajectories of atomic coordinates as simulated time progresses
- Researchers study trajectories by analyzing force profiles, energies, structural changes, etc.
- Visualization selections, graphics, structure properties recomputed for each trajectory timestep!



Cryo-EM / Cryo-ET Density Map Segmentation

Evaluate 3-D volumetric electron density maps and segment them, to identify key structural components

Index/label components so they can be referred to, colored, analyzed, and simulated...





Biomedical Technology Research Center for Macromolecular Modeling and Bioinformatics Beckman Institute, University of Illinois at Urbana-Champaign - www.ks.uiuc.edu

Density Map Segmentation



VMD GPU-accelerated density map segmentation of GroEL



Earnest, et al. J. Physical Chemistry B, 121(15): 3871-3881, 2017.



Biomedical Technology Research Center for Macromolecular Modeling and Bioinformatics Beckman Institute, University of Illinois at Urbana-Champaign - www.ks.uiuc.edu







NAMD on Summit, May 2018



NAMD simulations can generate up to 10TB of output per day on 20% of Summit APRIL 20, 2017 VOLUME 121 NUMBER 15 pubs.acs.org/JPCB





KLAUS SCHULTEN MEMORIAL ISSUE



Molecular Dyamics Trajectory Analysis

- MD simulations sample femtosecond timescales
- Millions of timesteps stored per trajectory
- Dynamics of biomolecular complexes are main interest, but solvent often accounts for half or more of the simulation content

Skip I/O for regions of bulk solvent where possible [1]

- Modern MD tools, e.g., VMD, NAMD, LAMMPS, HOOMD, employ extensive embedded scripting (Python, Tcl, etc) to permit simulation preparation, custom simulation protocols, analysis, and visualization
- Unified collective variables module allows identical analytical computations to be performed within LAMMPS, NAMD, and VMD, during pre-simulation modeling, in-situ, and post-hoc [2]

 [1] 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.
[2] Using collective variables to drive molecular dynamics simulations. G. Fiorin, M. L. Klein, and J. Hénin. Molecular Physics, 111:22-23, 3345-3362, 2013.

Petascale Molecular Dynamics I/O and Storage Challenges

- NAMD simulations can produce up to 10TB/day @ 1024 nodes (~20%) of ORNL Summit, more as optimizations raise NAMD performance further
- Petascale science campaigns require months of simulation runs
- Long-term storage of large-fractional petabytes impractical
- Historical "download output files for analysis and visualization" approach is a non-starter at this scale
- Demands visualization and analysis operate on the data in-place on the HPC system, whether post-hoc, in-transit, or in-situ
- Analyses must **identify salient features of structure, dynamics**, cull data that don't contribute to biomolecular processes of interest





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, CSCS Piz Daint, and similar systems



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



Biomedical Technology Research Center for Macromolecular Modeling and Bioinformatics Beckman Institute, University of Illinois at Urbana-Champaign - www.ks.uiuc.edu



VMD supports EGL for in-situ and parallel rendering on clouds, clusters, and supercomputers

- Eliminate dependency on windowing systems
- Simplified deployment of parallel VMD builds supporting off-screen rendering
- Maintains 100% of VMD OpenGL shaders and rendering features
- Support high-quality vendorsupported commercial OpenGL implementations in HPC systems that were previously limited to Mesa





Biomedical Technology Research Center for Macromolecular Modeling and Bioinformatics Beckman Institute, University of Illinois at Urbana-Champaign - www.ks.uiuc.edu

VMD EGL Rendering: Supports full VMD GLSL shading features Vulkan support coming soon...





Swine Flu A/H1N1 neuraminidase bound to Tamiflu

J. E.

High Performance Molecular Visualization: In-Situ and Parallel Rendering with EGL. J. E. Stone, P. Messmer, R. Sisneros, and K. Schulten. High Performance Data Analysis

and Visualization Workshop, IEEE IPDPSW, pp. 1014-1023, 2016.

64M atom HIV-1 capsid simulation



VMD EGL Performance on Amazon EC2 Cloud

MPI Ranks	EC2 "G2.8xlarge" GPU Instances	HIV-1 movie rendering time (sec), (I/O %) 3840x2160 resolution		
1	1	626s (10% I/O)		
2	1	347s (19% I/O)		
4	1	221s (31% I/O)		
8	2	141s (46% I/O)		
16	4	107s (64% I/O)		
32	8	90s (76% I/O)		

Performance at 32 nodes reaches ~48 frames per second

High Performance Molecular Visualization: In-Situ and Parallel Rendering with EGL. J. E. Stone, P. Messmer, R. Sisneros, and K. Schulten. High Performance Data Analysis and Visualization Workshop, IEEE IPDPSW, pp. 1014-1023, 2016.



64M atom HIV-1 capsid simulation rendered via EGL



NIH BTRC for Macromolecular Modeling and Bioinformatics http://www.ks.uiuc.edu/ Beckman Institute, U. Illinois at Urbana-Champaign

Next Generation: Simulating a Proto-Cell

- Emulate aspects of the *Mycoplasma mycoides* bacterium
- 200nm diameter
- ~1 billion atoms w/ solvent
- ~1400 proteins in membrane



Cryo-ET image of ultra-small bacteria (scale bar 100nm) Luef et al. Nature Comm., 6:6372, 2015.



Proto-Cell Data Challenges

- 1B-atom proto-cell requires nodes with more than TB RAM to build complete model...
- 1B-atom proto-cell binary structure file: 63GB
- Trajectory frame atomic coordinates: 12GB, 1.2TB/ns of simulation (1 frame per 10ps)
- Routine modeling and visualization tasks are a big challenge at this scale
 - Models contain thousands of atomic-detail components that must work together in harmony
 - Exploit persistent memory technologies to enable "instant on" operation on massive cell-scale models – eliminate several minutes of startup during analysis/visualization of known structure
 - Sparse output of results at multiple timescales will help ameliorate visualization and analysis I/O
 - Data quantization, compression, APIs like ZFP





High Fidelity Ray Tracing

- Interactive RT on laptops, desk, cloud ٠
- Large-scale parallel rendering: in situ or post hoc visualization tasks AO, DoF, instancing,
- Stereoscopic panorama and full-dome projections
- Omnidirectional VR: YouTube, HMDs
- Built-in ray tracing engines:
 - Tachyon: cross-platform RT
 - NVIDIA OptiX: GPU-accelerated and remote RT on VCA clusters
 - Intel OSPRay: CPU x86/Phi-optimized parallel rendering w/ MPI



VMD/OptiX all-atom Chromatophore

Lighting Comparison, STMV Capsid Two lights, no shadows Ambient occlusion + two









Satellite Tobacco Mosaic Virus Biomedical Technology Research Center for Macromolecular Modeling and Bioinformatics



Biomedical Technology Research Center for Macromolecular Modeling and Bioinformatics Beckman Institute, University of Illinois at Urbana-Champaign - www.ks.uiuc.edu



VMD w/ OptiX

- Interactive RT on laptops, desktops, and cloud
- Large-scale parallel rendering: in situ or post hoc visualization tasks
- Remote RT on NVIDIA VCA clusters
- Stereoscopic panoramic and full-dome projections
- Omnidirectional VR for YouTube, VR HMDs
- GPU memory sharing via NVLink

GPU-Accelerated Molecular Visualization on Petascale Supercomputing Platforms. J. E. Stone, K. L. Vandivort, and K. Schulten. UltraVis'13, pp. 6:1-6:8, 2013. Visualization of Energy Conversion Processes in a Light Harvesting Organelle at Atomic Detail. M. Sener, et al. SC'14 Visualization and Data Analytics Showcase, 2014. Chemical Visualization of Human Pathogens: the Retroviral Capsids. J. R. Perilla, B.-C. Goh, J. E. Stone, and K. Schulten. SC'15 Visualization and Data Analytics Showcase, 2015. Atomic Detail Visualization of Photosynthetic Membranes with GPU-Accelerated Ray Tracing. J. E. Stone et al., J. Parallel Computing, 55:17-27, 2016. Immersive Molecular Visualization with Omnidirectional Stereoscopic Ray Tracing and Remote Rendering J. E. Stone, W. R. Sherman, and K. Schulten. HPDAV, IPDPSW, pp. 1048-1057, 2016.



VMD/OptiX GPU Ray Tracing of all-atom Chromatophore w/ lipids.

HIV-1 Capsid



HIV-1 Parallel HD Movie Rendering on Blue Waters Cray XE6/XK7

New VMD TachyonL-OptiX on XK7 vs. Tachyon on XE6: K20X GPUs yield **up to twelve times** geom+ray tracing speedup

Ray Tracer Version	Node Type and Count	Script Load	State Load	Geometry + Ray Tracing	Total Time
New TachyonL-OptiX	64 XK7 Tesla K20X GPUs	2 s	39 s	435 s	476 s
New TachyonL-OptiX	128 XK7 Tesla K20X GPUs	3 s	62 s	230 s	295 s
TachyonL-OptiX [1]	64 XK7 Tesla K20X GPUs	2 s	38 s	655 s	695 s
TachyonL-OptiX [1]	128 XK7 Tesla K20X GPUs	4 s	74 s	331 s	410 s
TachyonL-OptiX [1]	256 XK7 Tesla K20X GPUs	7 s	110 s	171 s	288 s
Tachyon [1]	256 XE6 CPUs	7 s	160 s	1,374 s	1,541 s
Tachyon [1]	512 XE6 CPUs	13 s	211 s	808 s	1,032 s

 [1] GPU-Accelerated Molecular Visualization on Petascale Supercomputing Platforms.
J. E. Stone, K. L. Vandivort, and K. Schulten. UltraVis'13: Proceedings of the 8th International Workshop on Ultrascale Visualization, pp. 6:1-6:8, 2013.



20 M atom chromatophore patch



APRIL 20, 2017 **VOLUME 121** NUMBER 15 pubs.acs.org/JPCB





KLAUS SCHULTEN MEMORIAL ISSUE



www.acs.org



Volume: 169

Molecular Simulations and Visualization





nature THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE


Omnidirectional Stereoscopic Ray Tracing

- Ray trace 360° images and movies for Desk and VR HMDs: Oculus, Vive, Cardboard
- Stereo spheremaps or cubemaps allow very highframe-rate interactive OpenGL display
- AO lighting, depth of field, shadows, transparency, curved geometry, ...
- Summit 6x Tesla V100 GPU nodes:
 - Render many omni-stereo viewpoints, no acceleration structure rebuilds, tens of frames/sec per-node!
 - OptiX multi-GPU rendering, NVLink compositing and data distribution, etc...
 - Future: AI for warping between views

Atomic Detail Visualization of Photosynthetic Membranes with GPU-Accelerated Ray Tracing. J. E. Stone, et al. J. Parallel Computing, 55:17-27, 2016. Immersive Molecular Visualization with Omnidirectional Stereoscopic Ray Tracing and Remote Rendering. J. E. Stone, W. R. Sherman, and K. Schulten. High Performance Data Analysis and Visualization Workshop, IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), pp. 1048-1057, 2016.









Goal: Intuitive interactive viz. in crowded molecular complexes



Planetarium Dome Master Projections





Technology Opportunities and Collaborations

- Supercomputer Centers, Cray, IBM
 - Remote visualization
 - Performance, power profiling and optimization
- NVIDIA
 - GPU computing
 - Ray tracing
 - Remote visualization
 - ARM, Tablets, power profiling and optimization
- Intel
 - x86, Xeon Phi optimization
 - Ray tracing
- Amazon
 - Cloud deployment of VMD/NAMD, related tools
 - Remote visualization
- Universities:
 - G. Fiorin, J. Henin, Toni Giorgino, collective variables
 - T. Ertl, U. Stuttgart: visualization algorithms
 - M. Kuttel, U. Cape Town: visualization and analysis
 - W. Sherman, Indiana U.: VR HMDs, visualization

GPU computing, Ray tracing, Remote viz.



VR HMDs, 6DoF input devices



Energy efficiency: ARM+GPU

VMD is a Platform for Developing Research Tools Over 110 VMD Plugins, Half Developed by Users

- VMD user-extensible scripting w/ Tcl/Tk, Python
- User-developed plugins:
 - Alanine Scanning
 - Collective Variable Analyzer
 - Clustering Tool
 - Carbon Nanostructure Builder
 - TorsionPlot
 - RMSD Trajectory Tool
 - Many others...





Biomedical Technology Research Center for Macromolecular Modeling and Bioinformatics Beckman Institute, University of Illinois at Urbana-Champaign - www.ks.uiuc.edu

QwikMD: Guided MD Simulation and Training

Smooths initial learning curve (non-expert users)

Training: used in 4 Center workshops to-date

Speed up tedious simulation preparation tasks (expert users)

Reproducibility: detailed log of all steps

Interactive preparation, simulation, and analysis



Selected VMD Plugins: Center Developed, and User Developed

Analysis	Modeling	Visualization	Collaboration
APBSRun	AutoIonize	Clipping Plane Tool	Remote Control
CatDCD	AutoPSF	Clone Rep	Data Import and Plotting
Contact Map <u>GofRGUI</u>	Chirality	DemoMaster	Data Import
HeatMapper	Cionize	Dipole Watcher	Multiplot
ILSTools	Cispeptide	Intersurf	PDBTool
IRSpecGUI	CGTools	Navigate	MultiText
MultiSeq	Dowser	NavFly	Externally Hosted Plugins and
NAMD Energy NAMD Plot	ffTK	MultiMolAnim	Extensions
NetworkView	Inorganic Builder	Color Scale Bar	Check sidechains
NMWiz	MDFF	Remote	MultiMSMS
ParseFEP	Membrane	Palette Tool	Interactive Essential Dynamic
PBCTools	Merge Structs	ViewChangeRender	Mead Ionize
PMEpot PropKa GUI	Molefacture	ViewMaster	Clustering Tool
RamaPlot	Mutator	Virtual DNA Viewer	iTrajComp
RMSD Tool	Nanotube	VMD Movie Maker	Swap RMSD
RMSD Trajectory Tool	Psfgen	Simulation	Intervor
<u>RMSD Visualizer Tool</u>	<u>RESPTool</u>	AlaScan	<u>SurfVol</u>
Salt Bridges Sequence Viewer	RNAView	AutoIMD	vmdICE
Symmetry Tool	Solvate	IMDMenu	
Timeline	SSRestraints	NAMD GUI	
TorsionPlot	Topotools	NAMD Server	75 MolFile I/O Plugins:
VolMap	L	QMTool	structure, trajectory, sequence,

structure, trajectory, sequence, and density map

http://www.ks.uiuc.edu/Research/vmd/plugins/

Example VMD Visualization and Analysis Plugins



Bendix

Dahl ACE, Chavent M and Sansom MSP Bendix: intuitive helix geometry analysis and abstraction. *Bioinformatics* 2012 28(16): 2193-2194.





Normal Mode Wizard

Bakan A, Meireles LM, Bahar I ProDy: Protein Dynamics Inferred from Theory and Experiments. *Bioinformatics* 2011 27(11):1575-1577.

Making Our Research Tools Easily Accessible

- Cloud based deployment
 - Full virtual machines (known as "AMI" in Amazon terminology)
 - Amazon AWS EC2 GPU-accelerated instances: http://www.ks.uiuc.edu/Research/cloud/
- Docker "container" images available in NVIDIA NGC registry
 - Users obtain Docker images via registry, download and run on the laptop, workstation, cloud, or supercomputer of their choosing
 - https://ngc.nvidia.com/registry/
 - https://ngc.nvidia.com/registry/hpc-vmd

Molecular dynamics-based refinement and validation for sub-5 Å cryo-electron microscopy maps. Abhishek Singharoy, Ivan Teo, Ryan McGreevy, John E. Stone, Jianhua Zhao, and Klaus Schulten. *eLife*, 10.7554/eLife.16105, 2016. (66 pages).

QwikMD-integrative molecular dynamics toolkit for novices and experts. Joao V. Ribeiro, Rafael C. Bernardi, Till Rudack, John E. Stone, James C. Phillips, Peter L. Freddolino, and Klaus Schulten. *Scientific Reports*, 6:26536, 2016.

High performance molecular visualization: In-situ and parallel rendering with EGL. John E. Stone, Peter Messmer, Robert Sisneros, and Klaus Schulten. *2016 IEEE International Parallel and Distributed Processing Symposium Workshop (IPDPSW)*, pp. 1014-1023, 2016.



Clusters, Supercomputers

Workstations, Servers, Cloud

Molecular Dynamics Flexible Fitting (MDFF)



APS at Argonne





Electron microscopy

FEI microscope

ORNL Titan





Molecular dynamics-based refinement and validation for sub-5Å cryoelectron microscopy maps. A. Singharoy, I. Teo, R. McGreevy, J. E. Stone, J. Zhao, and K. Schulten. eLife 2016;10.7554/eLife.16105

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)

hexameric tubule



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)

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

Compute Pearson correlation to evaluate quality-of-fit between a reference cryo-EM density map and a **simulated density map** from an **all-atom structure**.





MDFF Cross Correlation TimelineRegions with poor fitRegions with good fit

MDFF Density Map Algorithm

- Build spatial acceleration data structures, optimize data for GPU
- Compute 3-D density map:

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

 Truncated Gaussian and spatial acceleration grid ensure linear time-complexity



3-D density map lattice point and the neighboring spatial acceleration cells it references

Single-Pass MDFF GPU Cross-Correlation



Parallel MDFF Cross Correlation Analysis on Cray XK7

Rabbit Hemorrhagic D	isease Virus (RHDV)	
Traj. frames	10,000	
Structure component selections	720	-
Single-node XK7 (projected)	336 hours (14 days)	
128-node XK7	3.2 hours 105x speedup	
2048-node XK7	19.5 minutes 1035x speedup	

Calculation of 7M CCs would take **5 years** using serial CPU algorithm!



Stone et al., Faraday Discuss., 169:265-283, 2014.

VMD Tesla V100 Cross Correlation Performance

Rabbit Hemorrhagic Disease Virus: 702K atoms, 6.5Å resolution Volta GPU architecture almost 2x faster than previous gen Pascal:

Application and Hardware platform	Runtime, Speed	lup vs. Chimera,	VMD+GPU
Chimera Xeon E5-2687W (2 socket) [1]	15.860s,	1x	
VMD-CUDA IBM Power8 + 1x Tesla K40 [2]	0.488s,	32x	0.9x
VMD-CUDA Intel Xeon E5-2687W + 1x Quadro K6000 [1,2]	0.458s,	35x	1.0x
VMD-CUDA Intel Xeon E5-2698v3 + 1x Tesla P100	0.090s,	176x	5.1x
VMD-CUDA IBM Power8 "Minsky" + 1x Tesla P100	0.080s,	198x	5.7x
VMD-CUDA Intel Xeon E5-2697Av4 + 1x Tesla V100	0.050s,	317x	9.2x
VMD-CUDA IBM Power9 "Newell" + 1x Tesla V100	0.049s,	323x	9.3x

[1] 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 Discussions 169:265-283, 2014.
 [2] Early Experiences Porting the NAMD and VMD Molecular Simulation and Analysis Software to GPU-Accelerated OpenPOWER Platforms. J. E. Stone, A.-P. Hynninen, J. C. Phillips, K. Schulten. International Workshop on OpenPOWER for HPC (IWOPH'16), LNCS 9945, pp. 188-206, 2016.

Clustering Analysis of Molecular Dynamics Trajectories: Requires I/O+Memory for All-Pairs of Trajectory Frames



GPU-Accelerated Molecular Dynamics Clustering Analysis with OpenACC. J.E. Stone, J.R. Perilla, C. K. Cassidy, and K. Schulten. In, Robert Farber, ed., Parallel Programming with OpenACC, Morgan Kaufmann, Chapter 11, pp. 215-240, 2016.



Use of Node-Local Burst Buffers and Non-Volatile Memory DIMMs

- Perform viz+analysis in-transit in node-local SSDs, persistent memory NVDIMMs
- ORNL Summit I/O:
 - Parallel FS: 2.5 TB/s
 - Node-local PCIe "burst buffer" SSDs: 10+ TB/sec, 7PB capacity
- Plenty of capacity for full-detail MD trajectories, could enable ~100x increase in temporal resolution in cases where it would be valuable to the science
- Enable all-pairs trajectory clustering analyses and resulting visualizations
- Future systems with NVDIMMs (3D Xpoint, phase change memory) could eventually provide bandwidths approaching DRAM
- Use NVDIMMs w/ mmap(), APIs like PMDK to perform formerly-out-of-core calculations using persistent memory:

https://github.com/pmem/pmdk

 Imagine future Summit-like machines w/ NVLink-connected GPUs w/ access to high-bandwidth persistent memory on each node





Trade FLOPS for Reduced I/O

ORNL Summit compute node:

- 6x Tesla V100 GPUs, 2x POWER9 CPUs
- GPUs Peak: ~46 DP TFLOPS, ~96 SP TFLOPS
- Peak IB rate per node: ~23GB/sec
- Ratio of FLOPS vs. I/O:

~2,000 DP FLOPS/byte, ~4000 SP FLOPS/byte

~16K FLOPS per FP word

Unconventional approach: Recompute to avoid I/O

Computing+Visualizing Molecular Orbitals

- Movies of simulation trajectories provides insight into results
- QM, and hybrid (QM/MM) MO visualizations historically done from huge "cube" files, impractical
- Store QM wavefunctions + Gaussian basis set, only 10s of KB per stored timestep compared to 100s of MB
- Recompute MO grid on-the-fly from QM basis set, huge decrease in RAM+I/O in exchange for heavy FP arithmetic

NAMD goes quantum: An integrative suite for hybrid simulations. Melo, M. C. R.; Bernardi, R. C.; Rudack T.; Scheurer, M.; Riplinger, C.; Phillips, J. C.; Maia, J. D. C.; Rocha, G. D.; Ribeiro, J. V.; Stone, J. E.; Neese, F.; Schulten, K.; Luthey-Schulten, Z.; Nature Methods, 2018.

http://dx.doi.org/10.1038/nmeth.4638

High Performance Computation and Interactive Display of Molecular Orbitals on GPUs and Multi-core CPUs. J. E. Stone, J. Saam, D. Hardy, K. Vandivort, W. Hwu, K. Schulten, *2nd Workshop on General-Purpose Computation on Graphics Processing Units (GPGPU-2), ACM International Conference Proceeding Series*, volume 383, pp. 9-18, 2009.



MO Kernel for One Grid Point (Naive C)

for (at=0; at <numatoms; at++)="" {<br="">int prim_counter = atom_basis[at]; <u>calc_distances_to_atom(&atompos[at], &xdist, &ydist, &zdist, &dist2, &xdiv);</u></numatoms;>	Loop over atoms
<pre>for (contracted_gto=0.0f, shell=0; shell < num_shells_per_atom[at]; shell++) { int shell type = shell symmetry[shell counter];</pre>	Loop over shells
<pre>for (prim=0; prim < num_prim_per_shell[shell_counter]; prim++) { float exponent = basis_array[prim_counter]; float contract_coeff = basis_array[prim_counter + 1]; contracted_gto += contract_coeff * expf(-exponent*dist2); prim_counter += 2; }</pre>	Loop over primitives: largest component of runtime, due to expf()
<pre>for (tmpshell=0.0f, j=0, zdp=1.0f; j<=shell_type; j++, zdp*=zdist) { int imax = shell_type - j; for (i=0, ydp=1.0f, xdp=pow(xdist, imax); i<=imax; i++, ydp*=ydist, xdp*=xdiv tmpshell += wave_f[ifunc++] * xdp * ydp * zdp; }</pre>	Loop over angular momenta (unrolled in real code)
<pre>value += tmpshell * contracted_gto; shell_counter++; }</pre>	

.

MO GPU Parallel Decomposition





Biomedical Technology Research Center for Macromolecular Modeling and Bioinformatics Beckman Institute, University of Illinois at Urbana-Champaign - www.ks.uiuc.edu



VMD C₆₀ MO Viz. Perf, 516x519x507 Grid: **@.13s/frame, avoids 3.8GB/s I/O** per-node



Hardware platform	Runtime,	Speedup	
IBM Power8 (ORNL 'crest') + 1x Tesla K40 [1]	3.49s,	1.0x	
Intel Xeon E5-2697Av4 + 1x Tesla V100	0.610s,	5.7x	
Intel Xeon E5-2697Av4 + 2x Tesla V100	0.294s,	11.8x	
Intel Xeon E5-2697Av4 + 3x Tesla V100	0.220s,	15.9x	
IBM Power9 "Newell" + 1x Tesla V100	0.394s,	8.8x	NVLink perf.
IBM Power9 "Newell" + 2x Tesla V100	0.207s,	16.8x	boost w/ no code tuning
IBM Power9 "Newell" + 3x Tesla V100	0.151s,	23.1x	(YET)
IBM Power9 "Newell" + 4x Tesla V100	0.130s,	26.8x	

[1] Early Experiences Porting the NAMD and VMD Molecular Simulation and Analysis Software to GPU-Accelerated OpenPOWER Platforms. J. E. Stone, A.-P. Hynninen, J. C. Phillips, K. Schulten. International Workshop on OpenPOWER for HPC (IWOPH'16), LNCS 9945, pp. 188-206, 2016.

MO Kernel Structure, Opportunity for NRTC JIT... Data-driven execution, but representative loop trip counts in (...)



Molecular Orbital Computation and Display Process Runtime Kernel Generation, NVRTC Just-In-Time (JIT) Compilation

One-time initialization

Initialize Pool of GPU Worker Threads Read QM simulation log file, trajectory

Preprocess MO coefficient data

eliminate duplicates, sort by type, etc...

Generate/compile basis set-specific CUDA kernel

For current frame and MO index, retrieve MO wavefunction coefficients

Compute 3-D grid of MO wavefunction amplitudes using basis set-specific CUDA kernel

Extract isosurface mesh from 3-D MO grid

Render the resulting surface

For each trj frame, for each MO shown

```
for (shell=0; shell < maxshell; shell++) {</pre>
```

float contracted_gto = 0.0f;

}

// Loop over the Gaussian primitives of CGTO
int maxprim = const_num_prim_per_shell[shell_counter];
int shell_type = const_shell_symmetry[shell_counter];
for (prim=0; prim < maxprim; prim++) {
 float exponent = const_basis_array[prim_counter];
 float contract_coeff = const_basis_array[prim_counter + 1];
 contracted_gto += contract_coeff * expf(-exponent*dist2);
 prim_counter += 2;</pre>

contracted_gto = 1.832937 * expf(-7.868272*dist2); contracted_gto += 1.405380 * expf(-1.881289*dist2); contracted_gto += 0.701383 * expf(-0.544249*dist2);



Runtime-generated dataspecific MO CUDA kernel compiled via CUDA NVRTC JIT...





AVX-512ER MO CGTO Loop

int maxprim = num_prim_per_shell[shell_counter];

```
int shelltype = shell_types[shell_counter];
```

for (prim=0; prim<maxprim; prim++) {</pre>

float exponent = basis_array[prim_counter];

float contract_coeff = basis_array[prim_counter + 1];

// contracted_gto += contract_coeff * exp(exponent*dist2);

__m512 expval = _mm512_mul_ps(_mm512_set1_ps(exponent * MLOG2EF), dist2);

// expf() approximation required, use (base-2) AVX-512ER instructions...

__m512 retval = _**mm512_exp2a23_ps(expval);**





 Γ_{40} at Γ_{40} and Γ_{40}



Performance of AVX-512ER Instrinsics vs. Autovectorization on KNL: Small 172x173x169 Grid

•	Intel C++ '15 autovectorization (fail):	220+ sec
•	Hand-coded SSE2 w/ existing thread scheme:	48.5 sec
•	Hand-coded AVX-512ER w/ existing thread scheme:	6.3 sec
•	Hand-coded AVX-512ER, refactoring thread pool:	0.2 sec
•	Hand-coded AVX-512ER tuned thread pool:	0.131 sec
•	Hand-coded AVX-512ER+FMA tweaks:	0.107 sec

Further improvement will require attention to details of cache behaviour and further tuning of low-level threading constructs for Xeon Phi/KNL



Acknowledgements

- Theoretical and Computational Biophysics Group, University of Illinois at Urbana-Champaign
- CUDA Center of Excellence, University of Illinois at Urbana-Champaign
- NVIDIA CUDA and OptiX teams
- Intel OSPRay team
- Funding:
 - DOE INCITE, ORNL Titan: DE-AC05-00OR22725
 - NSF Blue Waters: NSF OCI 07-25070, PRAC "The Computational Microscope", ACI-1238993, ACI-1440026
 - NIH support: 9P41GM104601, 5R01GM098243-02





"When I was a young man, my goal was to look with mathematical and computational means at the inside of cells, one atom at a time, to decipher how living systems work. That is what I strived for and I never deflected from this goal." – Klaus Schulten

Related Publications http://www.ks.uiuc.edu/Research/gpu/

- NAMD goes quantum: An integrative suite for hybrid simulations. Melo, M. C. R.; Bernardi, R. C.; Rudack T.; Scheurer, M.; Riplinger, C.; Phillips, J. C.; Maia, J. D. C.; Rocha, G. D.; Ribeiro, J. V.; Stone, J. E.; Neese, F.; Schulten, K.; Luthey-Schulten, Z.; Nature Methods, 2018. (In press)
- Challenges of Integrating Stochastic Dynamics and Cryo-electron Tomograms in Whole-cell Simulations.
 T. M. Earnest, R. Watanabe, J. E. Stone, J. Mahamid, W. Baumeister, E. Villa, and Z. Luthey-Schulten.
 J. Physical Chemistry B, 121(15): 3871-3881, 2017.
- Early Experiences Porting the NAMD and VMD Molecular Simulation and Analysis Software to GPU-Accelerated OpenPOWER Platforms. J. E. Stone, A.-P. Hynninen, J. C. Phillips, and K. Schulten. International Workshop on OpenPOWER for HPC (IWOPH'16), LNCS 9945, pp. 188-206, 2016.
- Immersive Molecular Visualization with Omnidirectional Stereoscopic Ray Tracing and Remote Rendering. J. E. Stone, W. R. Sherman, and K. Schulten. High Performance Data Analysis and Visualization Workshop, IEEE International Parallel and Distributed Processing Symposium Workshop (IPDPSW), pp. 1048-1057, 2016.
- **High Performance Molecular Visualization: In-Situ and Parallel Rendering with EGL.** J. E. Stone, P. Messmer, R. Sisneros, and K. Schulten. High Performance Data Analysis and Visualization Workshop, IEEE International Parallel and Distributed Processing Symposium Workshop (IPDPSW), pp. 1014-1023, 2016.
- Evaluation of Emerging Energy-Efficient Heterogeneous Computing Platforms for Biomolecular and Cellular Simulation Workloads. J. E. Stone, M. J. Hallock, J. C. Phillips, J. R. Peterson, Z. Luthey-Schulten, and K. Schulten.25th International Heterogeneity in Computing Workshop, IEEE International Parallel and Distributed Processing Symposium Workshop (IPDPSW), pp. 89-100, 2016.







- Atomic Detail Visualization of Photosynthetic Membranes with GPU-Accelerated Ray Tracing. J. E. Stone, M. Sener, K. L. Vandivort, A. Barragan, A. Singharoy, I. Teo, J. V. Ribeiro, B. Isralewitz, B. Liu, B.-C. Goh, J. C. Phillips, C. MacGregor-Chatwin, M. P. Johnson, L. F. Kourkoutis, C. Neil Hunter, and K. Schulten. J. Parallel Computing, 55:17-27, 2016.
- Chemical Visualization of Human Pathogens: the Retroviral Capsids. Juan R. Perilla, Boon Chong Goh, John E. Stone, and Klaus Schulten. SC'15 Visualization and Data Analytics Showcase, 2015.
- Visualization of Energy Conversion Processes in a Light Harvesting Organelle at Atomic Detail. M. Sener, J. E. Stone, A. Barragan, A. Singharoy, I. Teo, K. L. Vandivort, B. Isralewitz, B. Liu, B. Goh, J. C. Phillips, L. F. Kourkoutis, C. N. Hunter, and K. Schulten. Visualization and Data Analytics Showcase, 2014.
 ***Winner of the SC'14 Visualization and Data Analytics Showcase
- Runtime and Architecture Support for Efficient Data Exchange in Multi-Accelerator Applications. J. Cabezas, I. Gelado, J. E. Stone, N. Navarro, D. B. Kirk, and W. Hwu. IEEE Transactions on Parallel and Distributed Systems, 26(5):1405-1418, 2015.
- Unlocking the Full Potential of the Cray XK7 Accelerator. M. D. Klein and J. E. Stone. Cray Users Group, Lugano Switzerland, May 2014.
- 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 Discussions, 169:265-283, 2014.
- 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. Journal of Parallel Computing, 40:86-99, 2014.



- **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.



Related Publications http://www.ks.uiuc.edu/Research/gpu/

- 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 15th 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. E. 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.





- **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. E. 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.



