# ATPESC Track 4: 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)

3:45pm-4:30pm, St. Charles Amphitheater, Q Center,

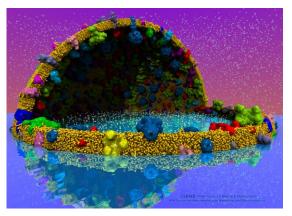
St. Charles, IL, Monday August 5th, 2019



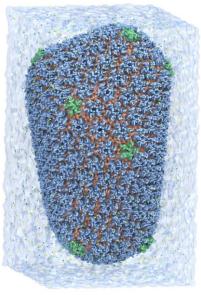


### VMD – "Visual Molecular Dynamics"

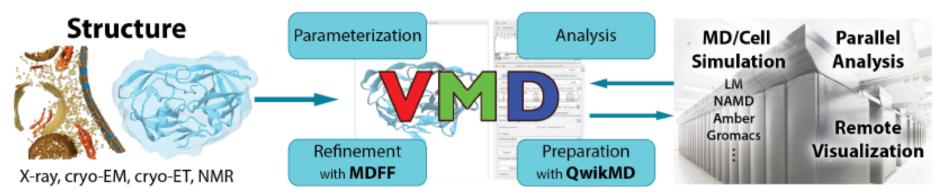
- 100,000 active users worldwide
- Visualization and analysis of:
  - Molecular dynamics simulations
  - Lattice cell simulations
  - Quantum chemistry calculations
  - Cryo-EM densities, volumetric data
- User extensible scripting and plugins
- http://www.ks.uiuc.edu/Research/vmd/



Cell-Scale Modeling



**MD** Simulation



### 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

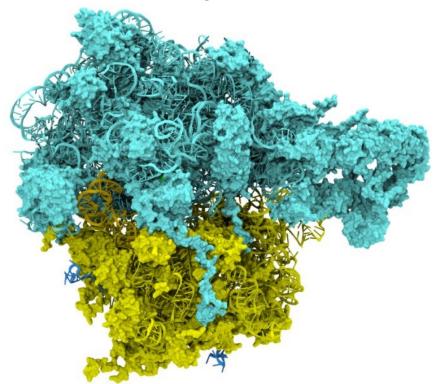




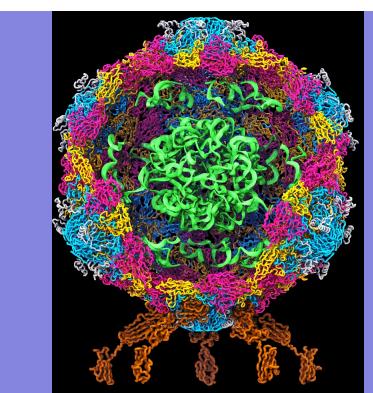
### Goal: A Computational Microscope

Study the molecular machines in living cells

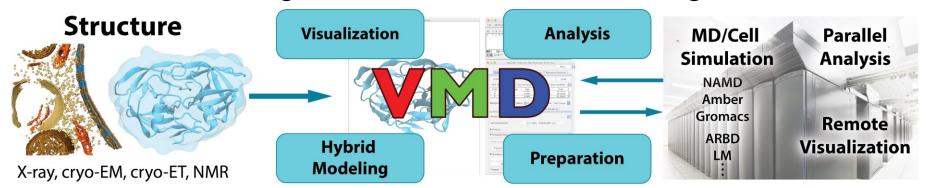
Ribosome: target for antibiotics



**Poliovirus** 



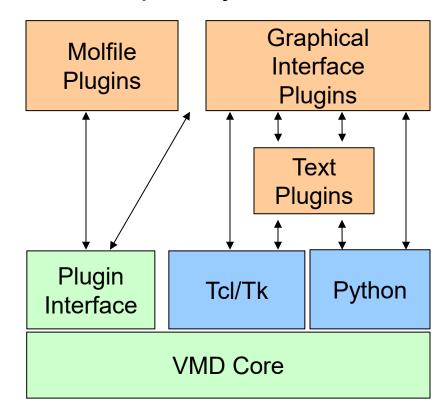
#### VMD: Building A Next Generation Modeling Platform



- Provide tools for simulation preparation, visualization, and analysis
  - Reach cell-scale modeling w/ all-atom MD, coarse grained, Lattice Microbes
  - Improved performance, visual fidelity, exploit advanced technologies (GPUs, VR HMDs)
- Enable hybrid modeling and computational electron microscopy
  - Load, filter, process, interpret, visualize multi-modal structural information
- Connect key software tools to enable state-of-the-art simulations
  - Support new data types, file formats, software interfaces
- Openness, extensibility, and interoperability are VMD hallmarks
  - Reusable algorithms made available in NAMD, for other tools

#### 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...



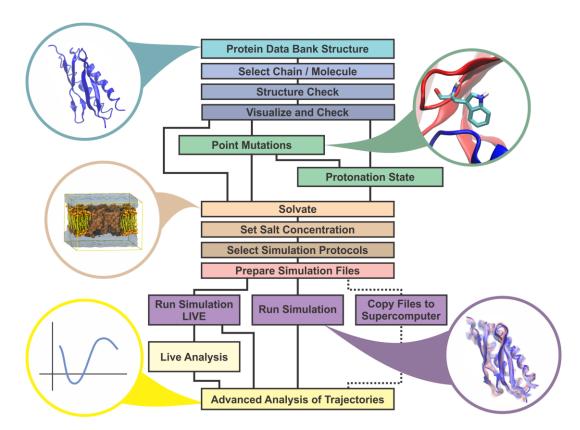
#### QwikMD: Guided MD Simulation and Training

Smooths initial learning curve (non-expert users)

Speed up tedious simulation preparation tasks (expert users)

### Reproducibility: detailed log of all steps

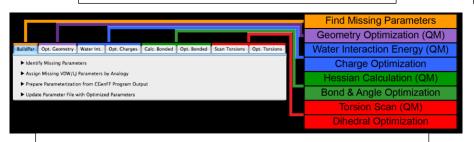
Interactive preparation, simulation, and analysis



### Parametrization with the Force Field Toolkit (FFTK)

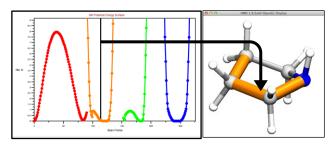
A plugin enabling the simulation of drug-like small molecules by lowering force field development barriers

A graphical interface organized to follow the standardized CGenFF workflow



Import data from the CGenFF Program webserver to assign initial parameters by analogy

#### Leverages VMD's powerful graphics capabilities to visualize parameter-related data within the context of molecular structure



Interactively explore energy profiles

#### **Future Plans:**

- Support free-for-academics QM software, ORCA (Neese, Max Planck)
- Parametrization tools for additional MM force fields
  - AMBER force field (Gumbart, Georgia Tech.)
  - Drude polarizable force fields (Roux, U. Chicago; MacKerell, U. Maryland)

#### Extensive training material available from the Center's website



**Documentation Website** 



Workshop Tutorial



**Screencast Demos** 

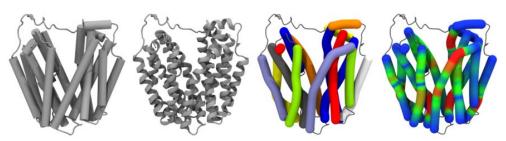
Selected VMD Plugins: Center Developed, and <u>User Developed</u>

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

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

structure, trajectory, sequence, and density map

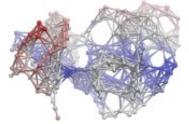
### **Example VMD Visualization and Analysis Plugins**



#### **Bendix**

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





#### **Normal Mode Wizard**

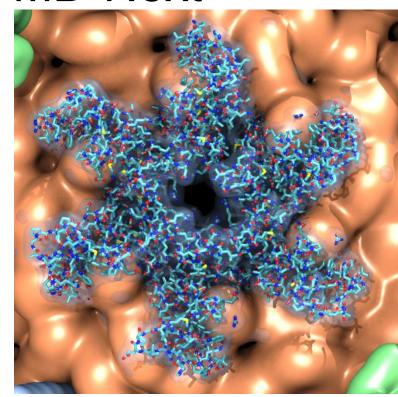
Bakan A, Meireles LM, Bahar I ProDy: Protein Dynamics Inferred from Theory and Experiments. *Bioinformatics* 2011 27(11):1575-1577.

## VMD 1.9.4 and VMD-Next

VMD "Coming Soon":

### VMD 1.9.4, and VMD-Next

- Python 3.x support
- New "molefacture" structure editor plugin
- Improved structure building and analysis tools
- High performance GPU structure+data clustering
- Density map and volume processing features: high performance GPU image segmentation, density map simulation, masking, visualization
- Many new and updated user-contributed plugins
- Deeper integration of interactive ray tracing
  - Seamless interactive RT in main VMD display window
  - Support trajectory playback in interactive RT
  - Enable multi-node interactive RT on HPC systems
- Built-in (basic) interactive remote visualization on HPC clusters and supercomputers



GPU Ray Tracing of HIV-1 Capsid Detail

### VMD Data Model

### VMD "Molecule"

- Collection of self-consistent molecular information
  - Atomic structure, time-varying atomic coordinates, velocities, ...
  - Volumetric data
  - Graphics objects
- Key missing information required for visualization is auto-generated heuristically – "guessed", if not provided in input files
  - Bond topology
- Limitations:
  - Fixed atom count per-timestep, requires use of "dummy atoms" for so-called "open system" simulations
  - VMD doesn't store or use MD force field parameters, various other data that are typically specific to particular simulation tools, force fields, etc.

### **Atomic Structure**

- Per-Atom attributes:
  - Name
  - Type
  - Residue ID, Unique residue ID, Residue Name, Residue Type
  - Bond list
  - Atomic number
  - Alternate location identifier
  - Insertion code
  - Chain
  - Segment name
  - Atom/Residue classification (protein, nucleic, hydrogen, water, "other")

## Other, Optional Per-Atom Fields

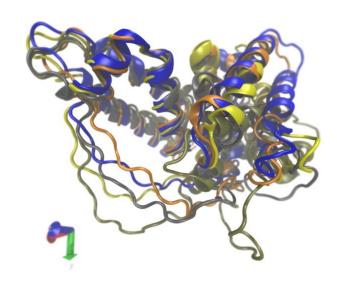
- Optional per-atom attributes:
  - Mass
  - Charge
  - Radius
  - Occupancy
  - B-factor
  - Fields primarily used by simulation preparation tools:
    - Bond Orders
    - Angles, Dihedrals, Impropers, Cross-Term Maps

## Time-Varying Trajectory Data

- Atomic coordinates
- Atomic velocities
- PBC unit cell
- QM, QM/MM molecular orbital data
- User-defined time-varying scalar quantities:
  - user, user2, user3, user4

# VMD Development Efforts Supporting Integrative Hybrid Modeling

- Extending mmCIF PDBx parser to encompass new IHM-specific records, data types
- Revising VMD "molfile plugin" APIs to communicate IHM data to VMD and represent it natively
- New atom selection keywords that encompass IHM structure data
- New graphical interfaces to query and interact with IHM data both quantitatively and visually



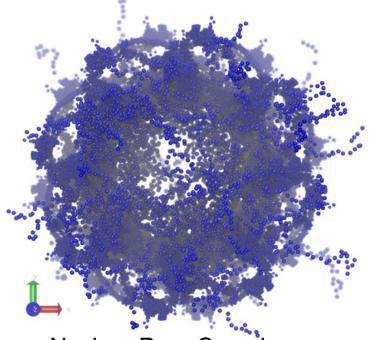
Serum Albumin Domains, PDB-DEV IHM #5





### Coarse-Grained IHM Data

- Coarse grained sphere/bead models
- Restraint information from experiments
- Multi-modal structure alignments, comparisons
- Linkage to underlying experimental images, statistics, etc.



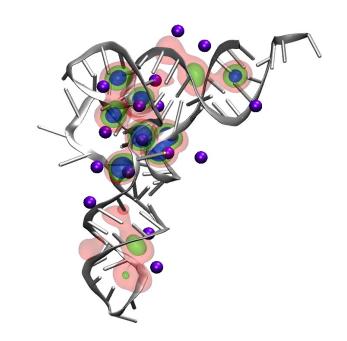
Nuclear Pore Complex, PDB-DEV IHM #12





### Display of Uncertainty, Error in IHM Models

- Query, visualize uncertainty, error, variance, in EM density maps, tomograms, atomic structure
- Requires IHM models to specify these statistics in the files
- Modeling tools, graphical interfaces can use this to guide user modeling tasks, analyses



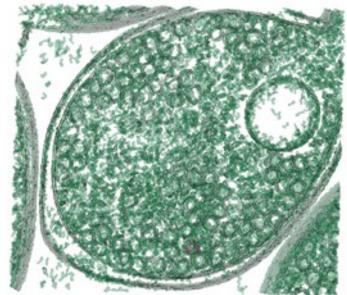
tRNA magnesium ion occupancy probability density surfaces, VMD volmap plugin





### Computing Challenges Posed by Large Hybrid Models

- Techniques like coarse graining allow modeling to reach the cell scale, but data sizes and interactivity remain a tremendous challenge
- Next-generation parallel- and GPUaccelerated computing techniques can make powerful analytical and visualization tools interactive for the first time:
  - Clustering analyses (structure RMSD, quality-of-fit, docking scores, etc)
  - Image segmentation, docking, alignment, fitting, coarse-graining...



VMD supports analysis and visualization of multi-gigavoxel EM tomograms, density maps



### Interoperability Challenges Posed by IHM Data

- Reliable interoperability between modeling and simulation tools hinges upon deep understanding of IHM data semantics
- When creating multi-modal hybrid models, quantitative and visual representation of "uncertainty" and "error" becomes important, so that the best information modalities are used to make analyses, decisions, and interpretations
- What are the physical properties assigned to coarse-grained "beads", Gaussians?
- Can those properties be understood among many different modeling tools?





# Overcoming the Biomolecular Orientation of VMD for Materials Science Work

- VMD molecular data structures are optimized for biomolecules
  - Expects atom/residue names that (roughly) follow RCSB Protein Databank conventions
  - Significant up-front structure analysis enables increased interactivity later
  - Cost of this is decreased generality in various cases:
    - Fixed atom count per timestep
    - Unrecognized atom naming conventions can inhibit use of some of the high-level atom selections
- Nanodevice and materials science data can be displayed in VMD
  - Assign per-atom name/type fields after loading
  - Enable VMD to better recognize structural elements
  - Example: xyz or LAMMPS dump files lack some expected fields, so it helps to use special features of LAMMPS plugin to assign them

### VMD LAMMPS Molfile Plugin

https://sites.google.com/site/akohlmey/software/lammps-plugin

- Reads LAMMPS text mode trajectory "dump" files
- Map LAMMPS per-atom fields to VMD molfile API:
  - LAMMPSREMAPFIELDS environment variable allows user to control mapping of peratom fields to VMD molfile API
  - List of mappings of LAMMPS fields into VMD fields, e.g. "VMD=LAMMPS"
  - Example to map LAMMPS forces into VMD velocity fields: set env(LAMMPSREMAPFIELDS) "vx=fx,vy=fy,vz=fz"
- Adapt LAMMPS trajectories with varying atom counts to VMD
  - LAMMPSMAXATOMS environment variable sets peak atom count for VMD, adding "dummy" atoms as placeholders:
    - Dummy atoms initialized with properties: name = @, type = X, resname = @@, segid = @@, chain = @, mass =0, charge=0, radius=0, element=X.
  - LAMMPSDUMMYPOS sets default atomic coords for dummy atoms that don't exist for a given frame:
    - set env(LAMMPSDUMMYPOS) {0.0 0.0 -10.0}

## **Exporting Molecular Data**

- VMD can be used to prepare MD simulations in conjunction with plugins and user scripting
- Need to watch out for assumptions made by heuristics, structure building tools
- Example: bond topology may need to be modified or deleted for some force fields
- Use the most appropriate structure manipulation tools for the MD package in use, e.g.:
  - NAMD: psfgen plugin
  - LAMMPS: topotools plugin

## VMD Visualization Concepts

### VMD Approach to Visualization

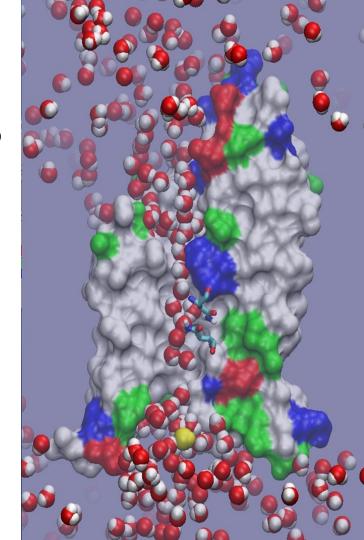
- Molecular scene is composed of "graphical representations"
- Each representation encapsulates a group of selected atoms, a drawing style, coloring style, and other parameters
- Representations are independent of each other, can be toggled on/off easily, allowing molecular scenes to be builtup incrementally
- 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.\*"

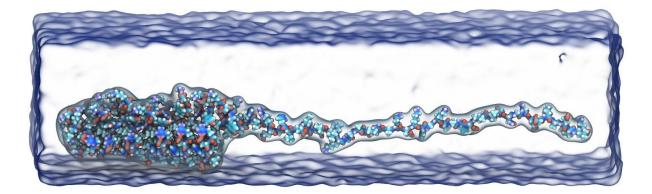
- Allows selection on user-defined data fields
- Promotes synergy between interactive and scripting interfaces, visualization and quantitative analysis tasks
- Works well with huge time-varying structures



### Structure Visualization

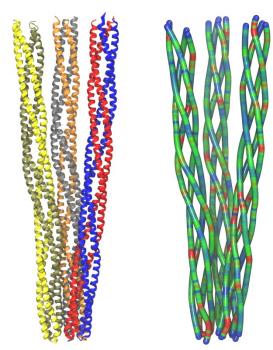
Molecular representations provide different levels of abstraction, atomic detail vs. higher level organizational information

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



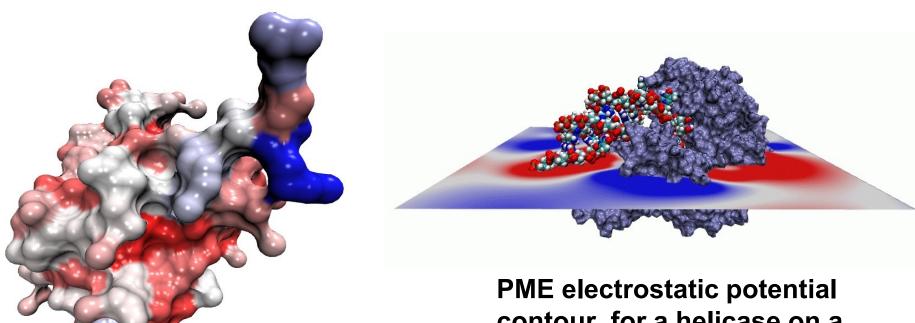
## Computed Properties

- 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

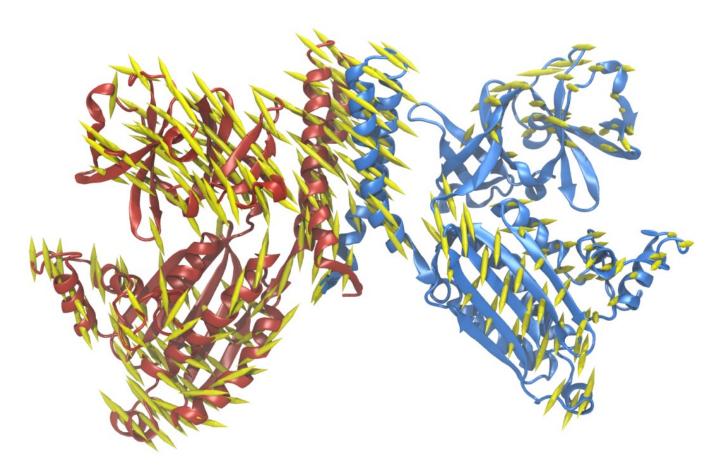
### Display of Computed Properties on Structures



Per-residue solvent-accessible surface area of Ubiquitin

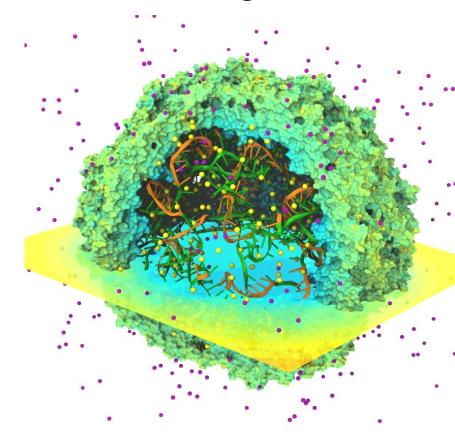
contour for a helicase on a volumetric slice plane

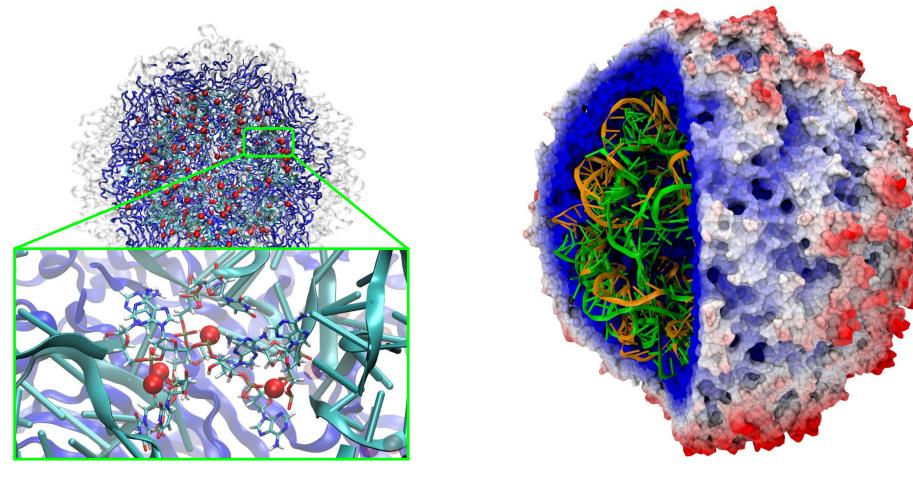
#### CheA kinase PCA: first principal component porcupine plot



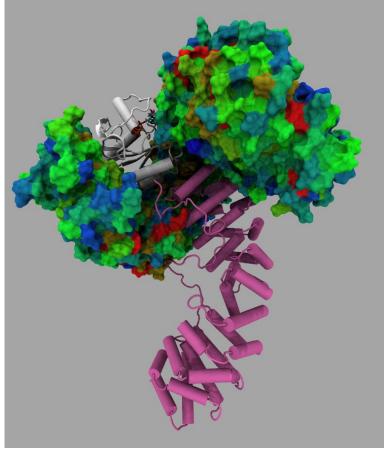
### Molecular Visualization Challenges

- Geometrically complex scenes
- Often show a mix of structural and spatial properties
- Time varying!
- Spatial relationships important to see clearly: depth cueing, shadows, AO helpful

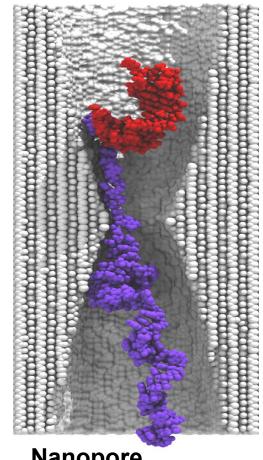




**Satellite Tobacco Mosaic Virus** 

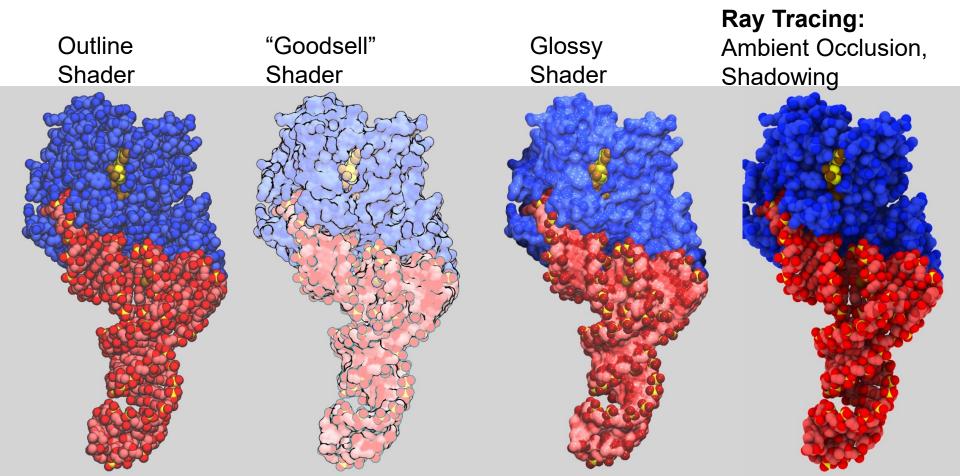


**Exportin Cse1p** 



Nanopore

### VMD Shading Comparison: EF-Tu



## VMD Scripting

## VMD Scripting

- Built-in Tcl interpreter is standard in all builds
- Custom builds can contain Python interpreter too
- Collection of simple VMD scripts:

http://www.ks.uiuc.edu/Research/vmd/script\_library/

### **Atom Selections**

```
set ionsel [atomselect top "ions"]
set watsel [atomselect top "water"]
set solute [atomselect top "not water and not ions"]
```

```
set numions [$ionsel num]
set coords [$solute get {x y z}]
```

### Built-in "Measure" Analysis Routines

- Provide high-performance routines for common analysis functions
- Operate on atom selections, sometimes over a frame range set com [measure center \$sel weight mass] set avpos [measure avpos \$sel first 0 last 999 step 1] set c [measure contacts 1.0 \$sel1 \$sel2]

```
set txmat [measure fit $sel1 $sel2]
$sel2 move $txmat
```

```
# Print RMSD between two selections between the first timestep and each
# later timestep for the given molecule id (default: top)
proc print rmsd trajectory {{mol top}} {
                                                         # use frame 0 for the reference
  set reference [atomselect $mol "protein" frame 0]
  set compare [atomselect $mol "protein"]
                                                         # the frame being compared
  set numsteps [molinfo $mol get numframes]
                                                         # query number of trajectory frames
  for {set frame 0} {$frame < $numsteps} {incr frame} {</pre>
                                                         # set the correct comparison frame
     $compare frame $frame
     set transmat [measure fit $compare $reference]
                                                         # compute the transformation
     $compare move $transmat
                                                         # align comparison to reference
     set rmsd [measure rmsd $compare $reference]
                                                         # compute the RMSD
     puts "RMSD of $frame is $rmsd"
                                                         # print the RMSD
```

Trajectory Analysis and

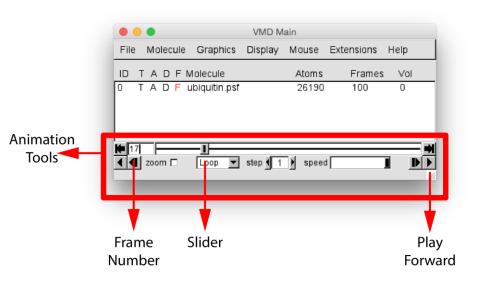
Visualization

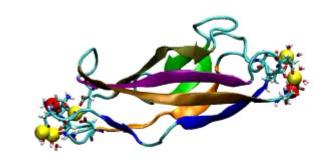
#### Molecular Dyamics Trajectory Analysis

- 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.

### Visualization of MD Trajectories

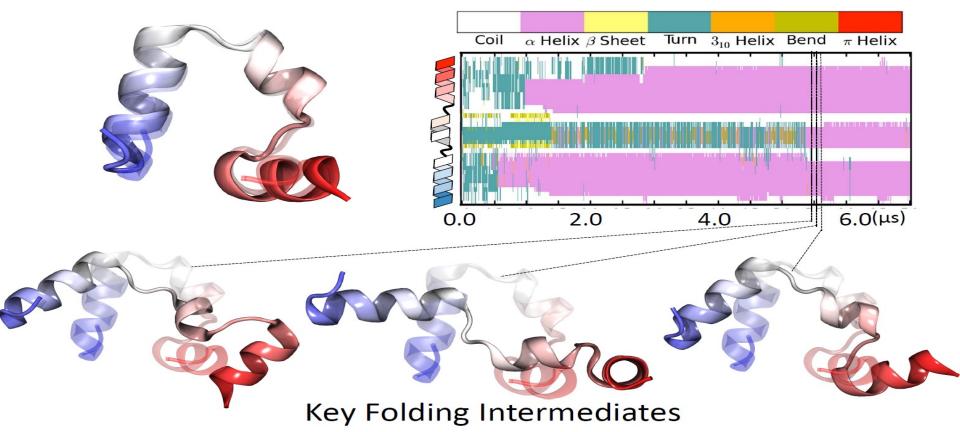
- Allow researchers study trajectories by analyzing force profiles, energies, structural changes, etc.
- Visualization selections, graphics, structure properties, colors can all be recomputed for each trajectory timestep!





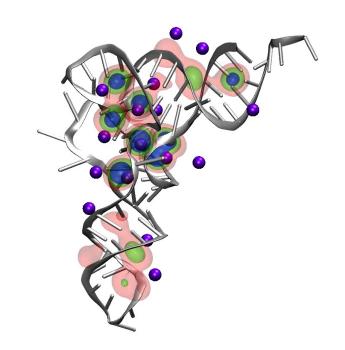
# Folding Dynamics of Villin Headpiece Unveiled 6.9µs folding simulation of 30K atoms: 380GB trajectory

Schulten et al. *Biophys J* **94**:**L**75, 2008, **97**: 2009



#### Time-Averaged Volumetric Properties

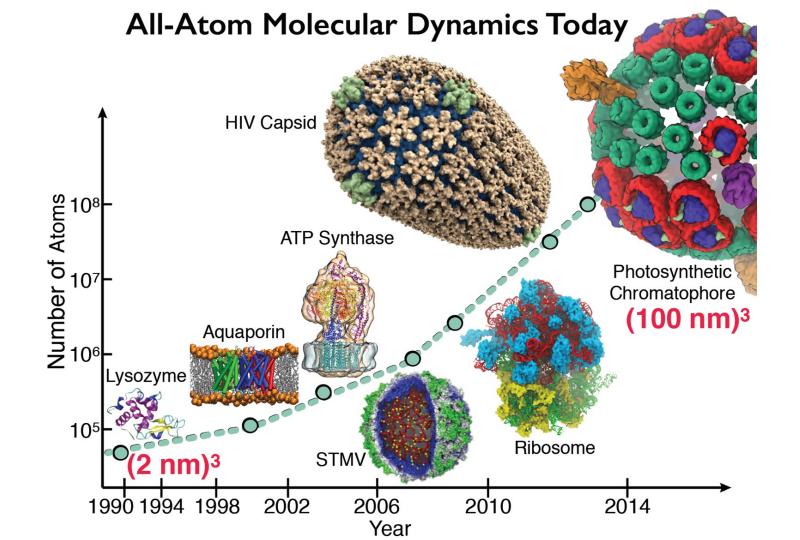
- Compute density, distance, occupancy, potential maps for a frame or averaged over a trajectory
- Example: display binding sites for diffusively bound ions as probability density isosurfaces



tRNA magnesium ion occupancy: VMD volmap plugin

Large System Analysis and

Visualization







#### VMD Petascale Visualization and Analysis

- Combination of growing system sizes and timescales of simulation trajectories poses a major data size challenge for molecular visualization and analysis
- Parallel I/O rates up to 275 GB/sec on 8192 Cray
   XE6 nodes can read in 231 TB in 15 minutes!
- Analyze/visualize large trajectories too large to transfer off-site:
  - User-defined parallel analysis operations, data types
  - Parallel rendering, movie making
- Supports GPU-accelerated compute nodes for both visualization and analysis tasks:
  - GPU accelerated trajectory analysis w/ CUDA
  - OpenGL and GPU ray tracing for visualization and movie rendering

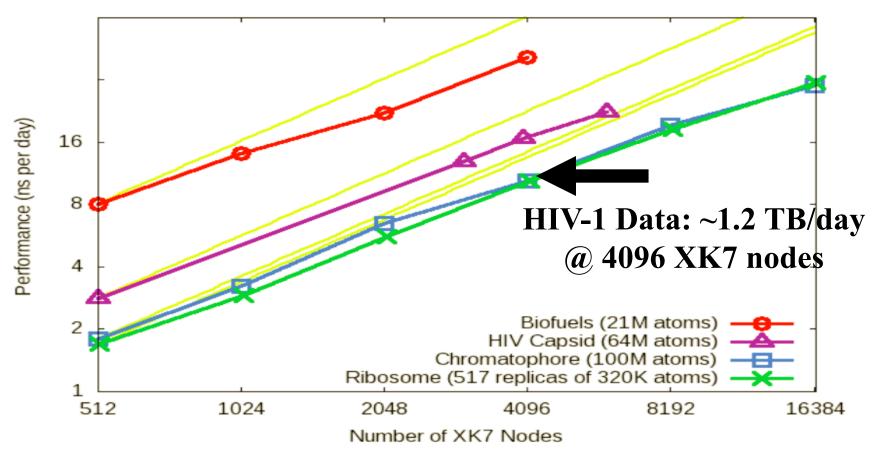


NCSA Blue Waters Hybrid Cray XE6 / XK7 22,640 XE6 dual-Opteron CPU nodes 4,224 XK7 nodes w/ Telsa K20X GPUs

Parallel VMD currently available on: ORNL Summit and Titan, NCSA Blue Waters, IU Big Red II, CSCS Piz Daint, many similar systems

#### NAMD Titan XK7 Performance August 2013

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



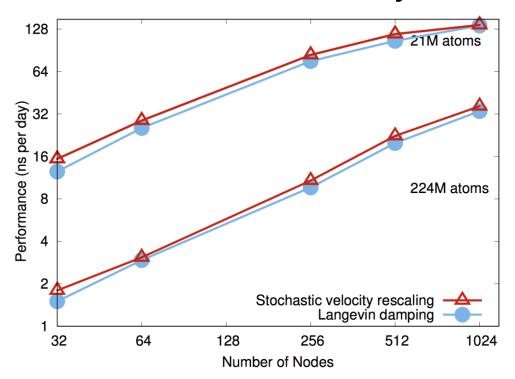
# Petascale Molecular Dynamics I/O and Storage Challenges

- 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





#### NAMD on Summit, May 2018

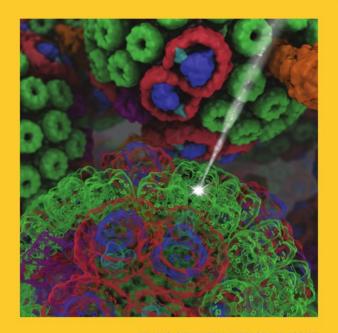


NAMD simulations can generate up to 10TB of output per day on 20% of Summit

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

# THE JOURNAL OF PHYSICAL CHEMISTRY



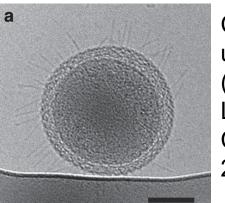


KLAUS SCHULTEN MEMORIAL ISSUE

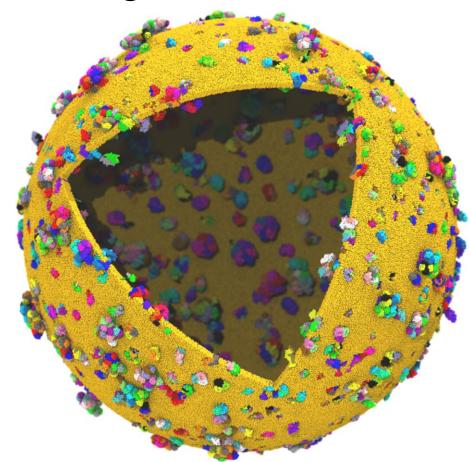


#### Next Generation: Simulating a Proto-Cell

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

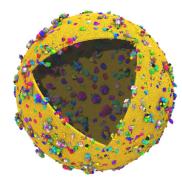


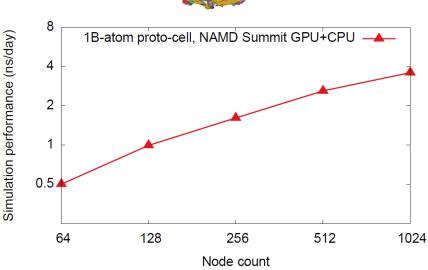
Cryo-ET image of ultra-small bacteria (scale bar 100nm) Luef et al. Nature Comm., 6:6372, 2015.



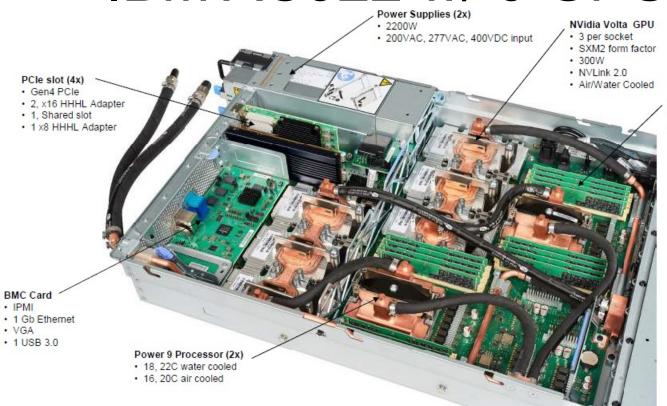
#### Proto-Cell Data Challenges

- 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





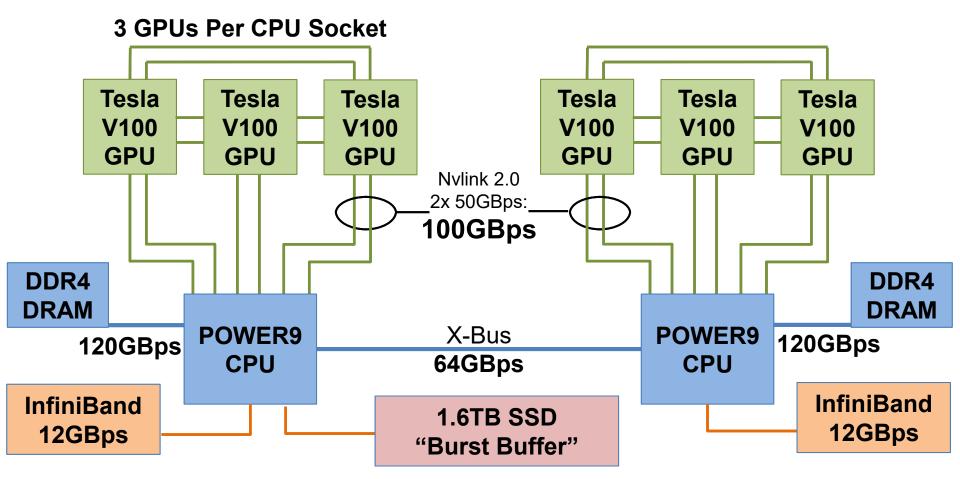
## IBM AC922 w/ 6 GPUs



#### Memory DIMM's (16x)

- · 8 DDR4 IS DIMMs per socket
- · 8, 16, 32,64, 128GB DIMMs

#### IBM AC922 Summit Node

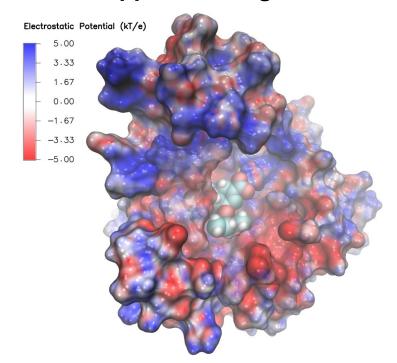


VMD Off-Screen Rendering w/ EGL

- Containers+Cloud+Workstations with recent NVIDIA drivers
- VMD on HPC systems w/ latest GPUs:
  - Cray XC50, CSCS Piz Daint
  - ORNL Summit in progress now
  - IBM OpenPOWER, drivers 375.66 and later support both GLX and EGL



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



Swine Flu A/H1N1 neuraminidase bound to Tamiflu

64M atom HIV-1 capsid simulation

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.

# Running VMD on Distributed Memory Clouds, Clusters, and Supercomputers with MPI

## Using VMD MPI Builds

- See "Running VMD on Supercomputers" section in the VMD tutorial
- Run one MPI rank (VMD process) per compute node
  - Each rank uses all CPU cores and all GPU accelerators they find
  - GPUs are shared between graphics/visualization and computing tasks
- Parallel VMD Scripting APIs:
  - "parallel" script subcommands, e.g., "parallel for ...":
    - "nodename", "noderank", "nodecount", "barrier", "allgather", "allreduce", "for"
  - Available even in non-MPI builds, for single-node script compatibility
- Work scheduling:
  - Intra-node work scheduled via multithreading, CUDA, etc.
  - Inter-node distributed memory work scheduling of user analysis/viz scripts is provided by built-in load balancing implementation for VMD





#### Example MPI Session, CSCS Piz Daint

stonej@daint103> srun -C gpu -n 256 --ntasks-per-node=1 \
/users/stonej/local/bin/vmd193 -dispdev text -e rendermovie.tcl

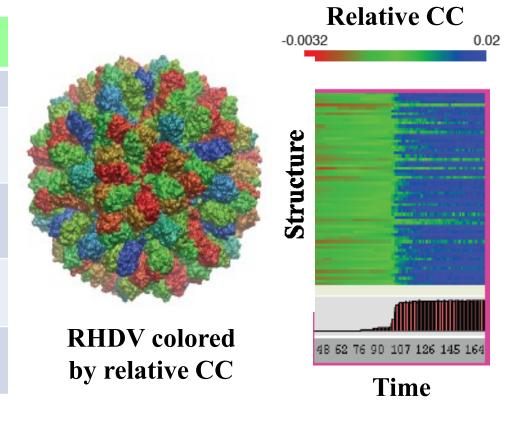
- Info) VMD for CRAY XC, version 1.9.3 (December 15, 2016)
- Info) http://www.ks.uiuc.edu/Research/vmd/
- Info) Email questions and bug reports to vmd@ks.uiuc.edu
- Info) Please include this reference in published work using VMD:
- Info) Humphrey, W., Dalke, A. and Schulten, K., 'VMD Visual
- Info) Molecular Dynamics', J. Molec. Graphics 1996, 14.1, 33-38.
- Info) -----
- Info) Creating CUDA device pool and initializing hardware...
- Info) Initializing parallel VMD instances via MPI...
- Info) Found 256 VMD MPI nodes containing a total of 6144 CPUs and 256 GPUs:
- Info) 0: 24 CPUs, 60.8GB (96%) free mem, 1 GPUs, Name: nid03072
- Info) 1: 24 CPUs, 60.8GB (96%) free mem, 1 GPUs, Name: nid03073
- [...example output omitted...]
- Info) 254: 24 CPUs, 60.9GB (96%) free mem, 1 GPUs, Name: nid03376
- Info) 255: 24 CPUs, 60.9GB (96%) free mem, 1 GPUs, Name: nid03377





#### Parallel MDFF Cross Correlation Analysis on Cray XK7

Rabbit Hemorrhagic Disease 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!

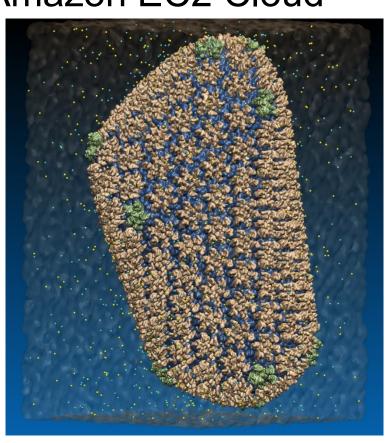
Stone et al., Faraday Discuss., 169:265-283, 2014.

#### VMD 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 FPS

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



64M atom HIV-1 capsid simulation rendered via EGL

# 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.

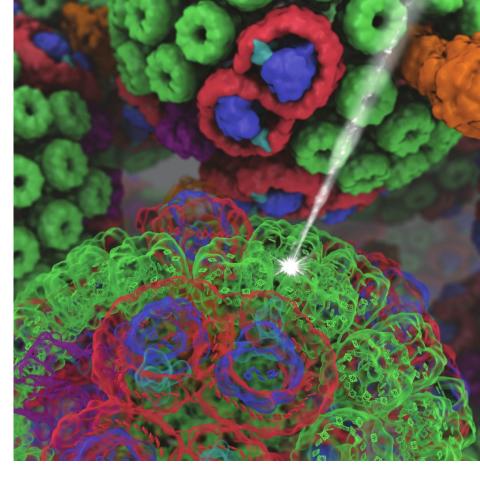
J. E. Stone, K. L. Vandivort, and K. Schulten. UltraVis'13: Proceedings of the 8th International Workshop on Ultrascale Visualization, pp. 6:1-6:8, 2013.

High Fidelity Ray Tracing for Interactive

and Cinematic Visualization

#### High Fidelity Ray Tracing

- Advanced rendering techniques save scientists time, produce images that are easier to interpret
- Ambient Occlusion, Depth of Field, high quality transparency, instancing, ....
   Interactive RT on laptops, desk, cloud,
- Interactive RT on laptops, desk, cloud and remote supercomputers
- Large-scale MPI parallel rendering: in situ or post hoc visualization tasks
- Stereoscopic panorama and full-dome projections
- Omnidirectional VR: YouTube, HMDs
- Built-in ray tracing engines:
  - Tachyon: cross-platform RT
  - NVIDIA OptiX: GPU RTX-accelerated
  - Intel OSPRay: CPU x86-optimized

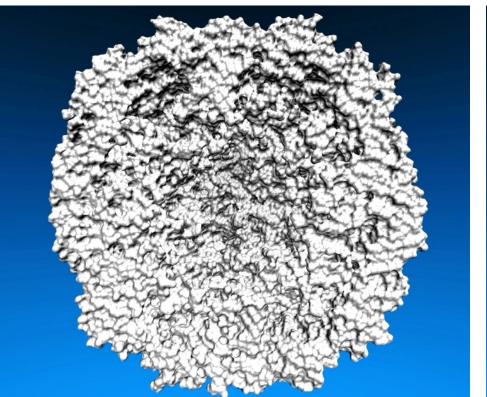


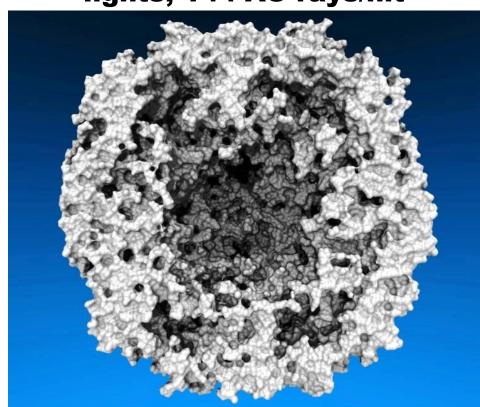
VMD/OptiX all-atom Chromatophore

#### Lighting Comparison, STMV Capsid

Two lights, no shadows

Ambient occlusion + two lights, 144 AO rays/hit

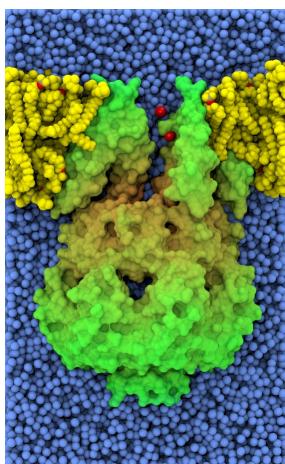




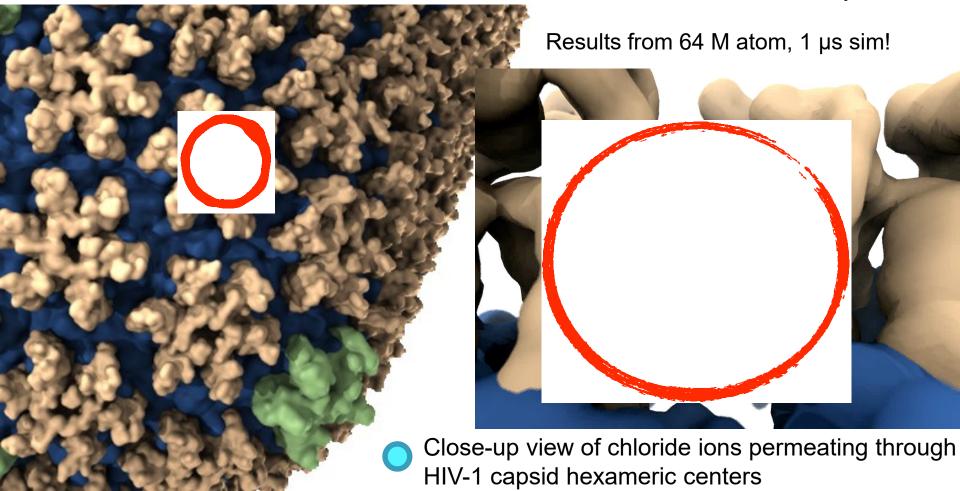
#### Geometrically Complex Scenes

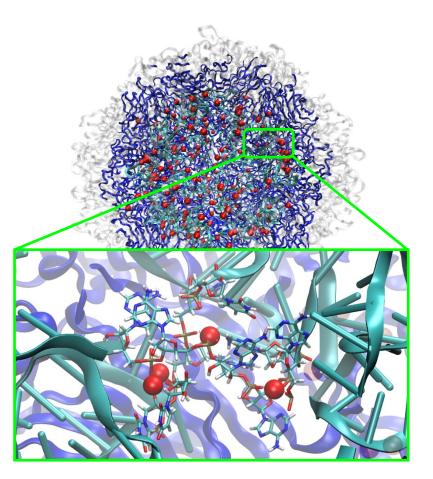
# Ray tracing techniques well matched to molecular viz. needs:

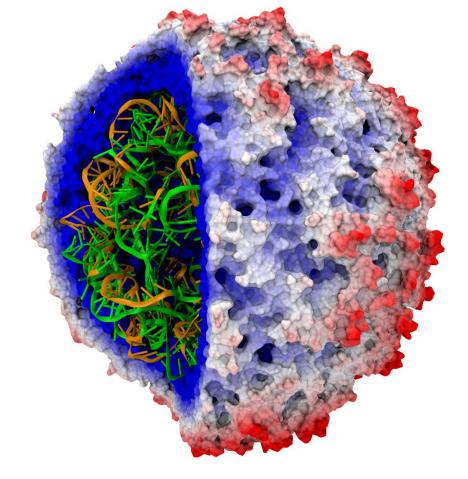
- Curved geometry, e.g. spheres, cylinders, toroidal patches, easily supported
- Greatly reduced memory footprint vs. polygonalization
- Runtime scales only moderately with increasing geometric complexity
- Occlusion culling is "free", RT acceleration algorithms do this and much more



#### Goal: Intuitive interactive viz. in crowded molecular complexes





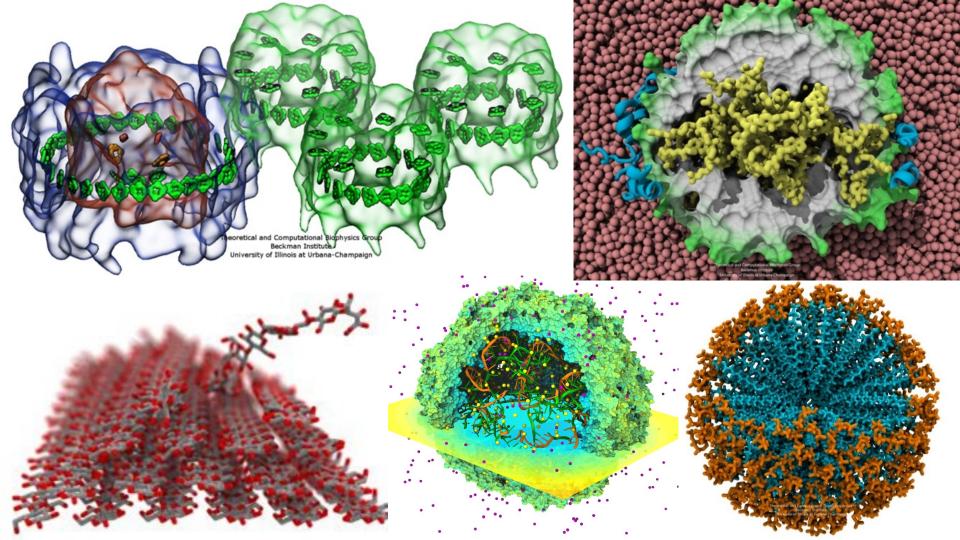


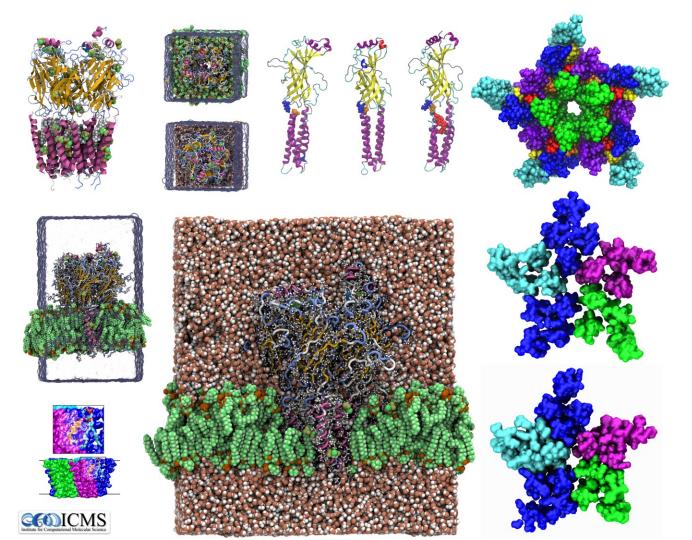
#### Satellite Tobacco Mosaic Virus Biomedical Technology Research Center for Macromolecular Modeling and Bioinformatics

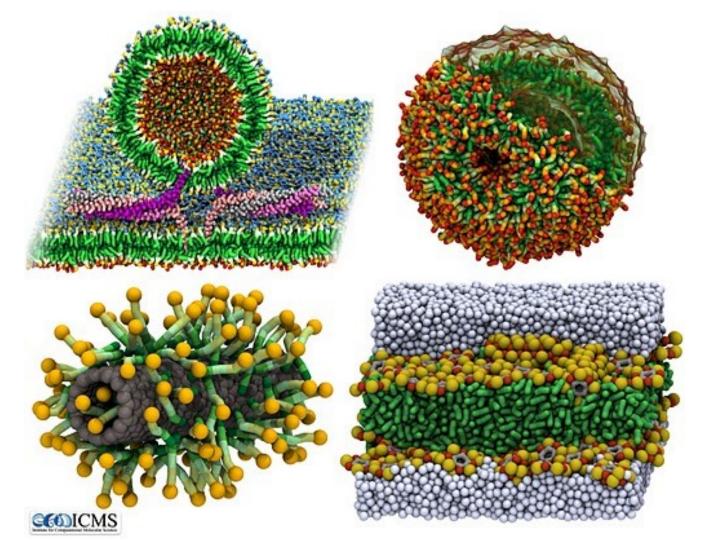
Beckman Institute, University of Illinois at Urbana-Champaign - www.ks.uiuc.edu

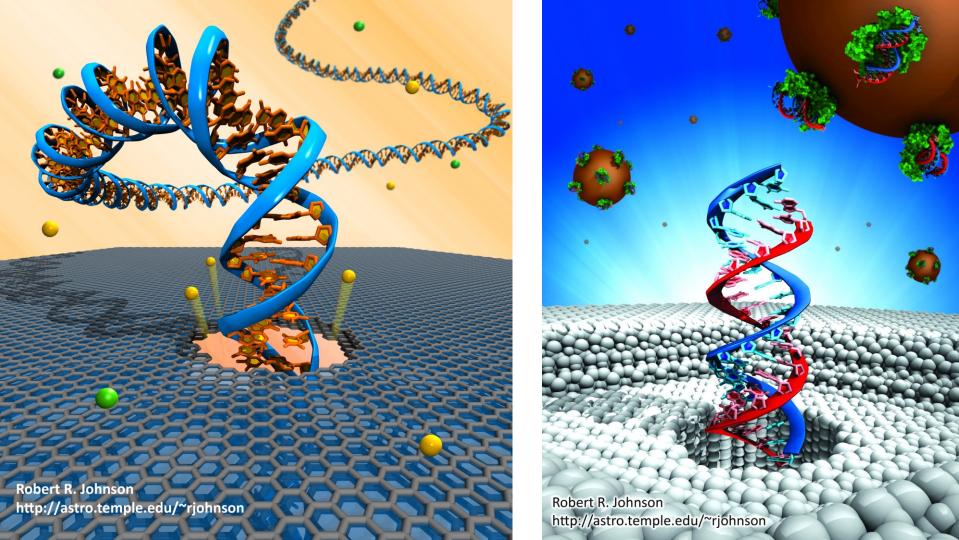


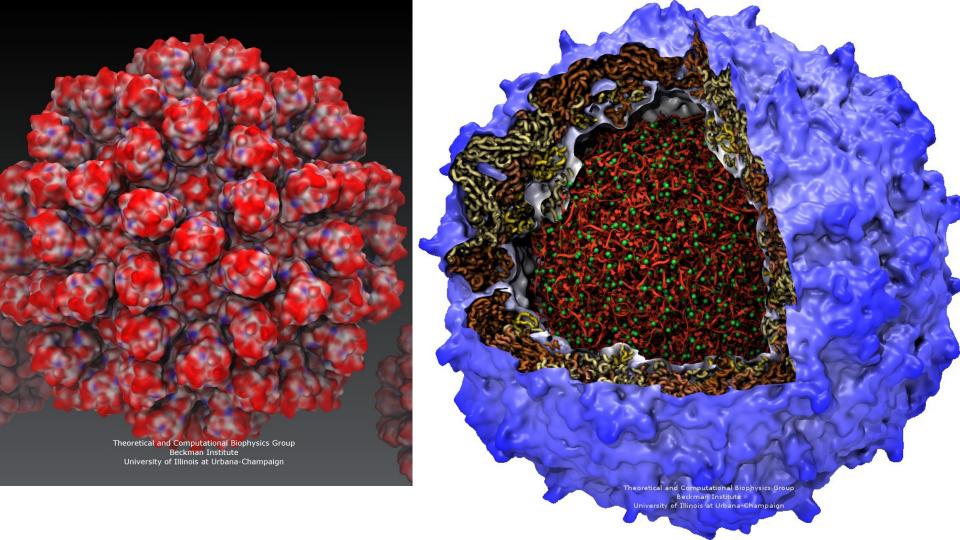












### VMD w/ OptiX 6

- Interactive RT on laptops, desktops, and cloud
- Large-scale parallel rendering: in situ or post hoc visualization
- Remote ray tracing with NvPipe video streaming
- Stereoscopic panoramic and full-dome projections
- Omnidirectional VR for YouTube, VR HMDs
- VMD+OptiX NGC container: https://ngc.nvidia.com/registry/
- GPU memory sharing via NVLink
- In-progress:

Denoising: faster turnaround w/ AO, DoF, etc

**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.

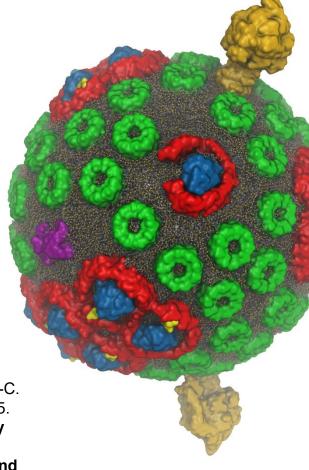
Visualization of Energy Conversion Processes in a Light Harvesting Organelle at Atomic Detail. M. Sener, et al. SC'14 Visualization and Data Analytics Showcase, 2014.

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

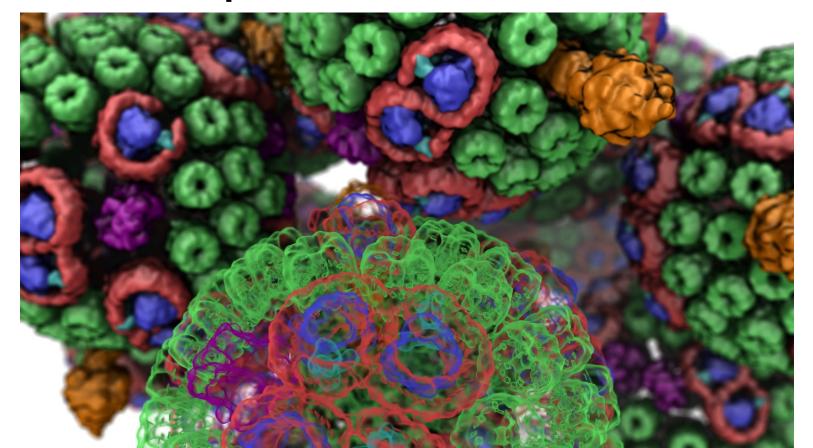
**Tracing.** J. E. Stone et al., J. Parallel Computing, 55:17-27, 2016.

Immersive Molecular Visualization with Omnidirectional Stereoscopic Ray Tracing and Remote Rendering J. E. Stone, W. R. Sherman, and K. HPDAV, IPDPSW, pp. 1048-1057, 2016.



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

### VMD/OpiX RTX Acceleration







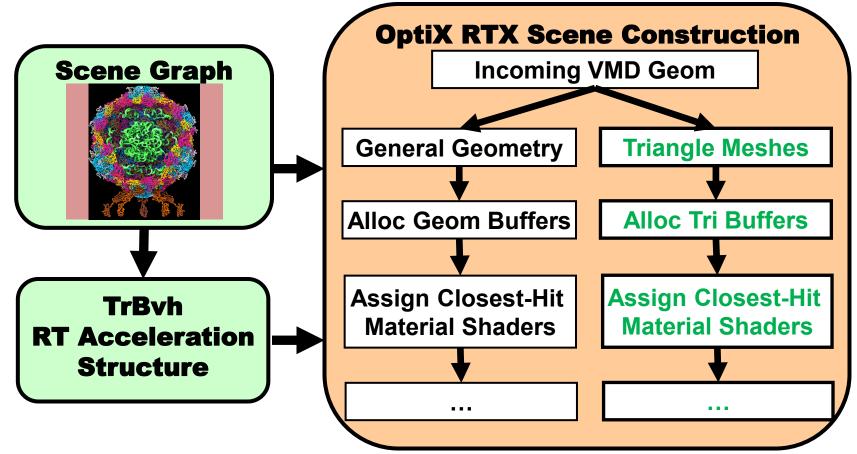
#### What does RTX Acceleration do for VMD?

- Hardware acceleration (Turing RT cores) of:
  - BVH AS traversal
  - Ray-triangle intersection
- BVH AS can embed triangle geometry
  - Triangle geometry buffer can be ephemeral
- Often gives VMD a ~6x performance gain



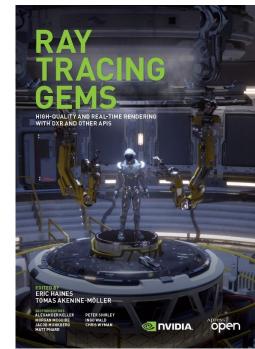


#### VMD Scene w/ OptiX RTX APIs



## Secondary Benefits and RTX Performance Observations

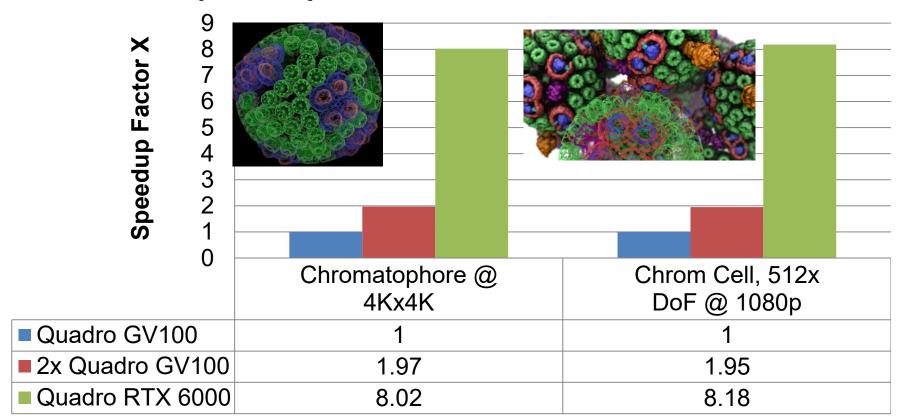
- Traversal and intersection work performed by RT cores vacates GPU SMs and makes them available for other RT work
- RTX hardware traversal performance approaches GPU memory bandwidth limits
- Future RTX hardware could end up being bandwidth bound in some cases
- Start adapting geometric data representations for minimum footprint, e.g. by using compressed or quantized data representations such as Octahedron Normal Vector encoding (replace 3x 32-bit floats with a single 32-bit int)







#### VMD OptiX RT performance on Quadro RTX 6000

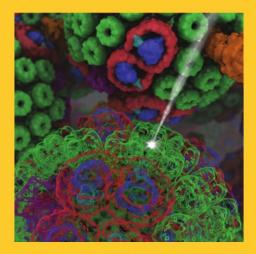




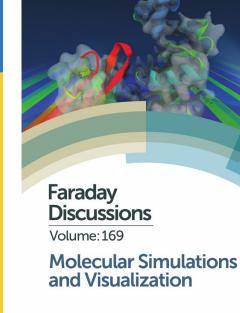


APRIL 20, 2017
VOLUME 121
NUMBER 15
pubs.acs.org/JPCB

## THE JOURNAL OF PHYSICAL CHEMISTRY

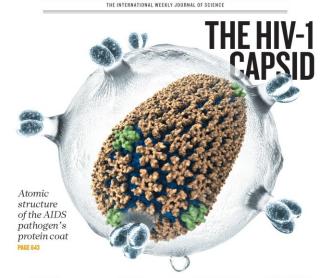


KLAUS SCHULTEN MEMORIAL ISSUE









THE FIRST LIGHT In pursuit of the most distant galaxies PAGE 554 CROSSING THE BORDERS International collaborations make the most impact

A SITTING TARGET An indirect hit on 'undruggable' KRAS protein





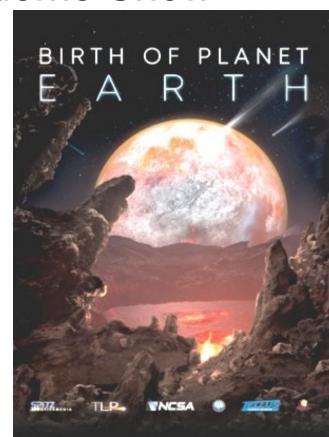
www.acs.org

VMD Application Examples and

**Implementation Details** 

# Cinematic Rendering: "Birth of Planet Earth" Fulldome Show

- Joint project with:
  - NCSA Advanced Visualization Lab
  - Thomas Lucas Productions, Inc.
  - Spitz Creative Media
  - Tellus Science Museum
- NSF Support: CADENS award ACI-1445176
- Premiered March 28, Zeiss Großplanetarium, Berlin, Germany



### Why Use VMD for Cinematic Rendering

- Use all science materials where possible
- Maintain peak efficiency for calculating molecular geometries and representations
- OptiX-based RT engine and data structures designed for memory efficiency and performance specifically for molecular models
- Built-in scripting engines enabled control and incorporation within cinematic pipeline
- Avoid I/O bottlenecks



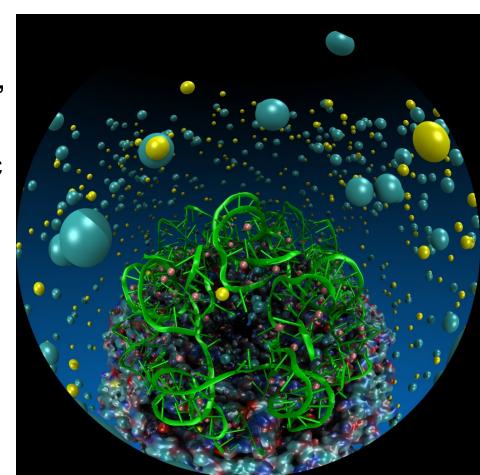


#### VMD Planetarium Dome Master Camera

- Fully interactive RT with ambient occlusion, shadows, depth of field, reflections, ...
- Both mono and stereoscopic
- No post-processing required

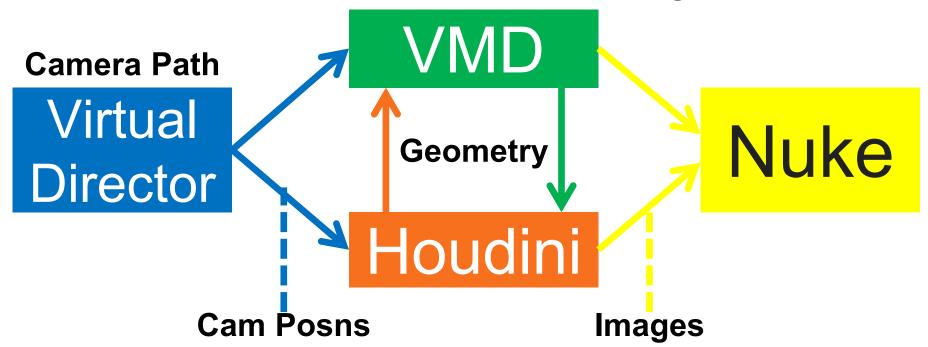


Ray Tracing Gems Ch. 4



### "Birth of Planet Earth" Cinematic Pipeline

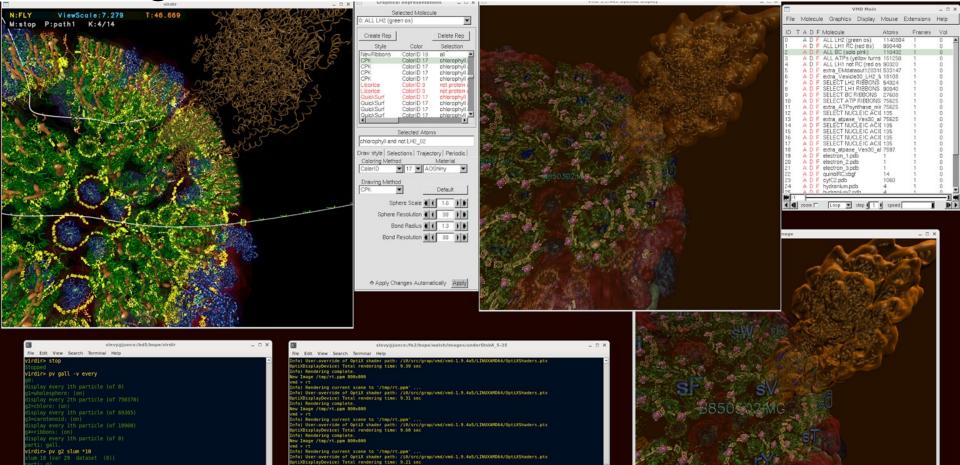
**Visualization and Rendering** 

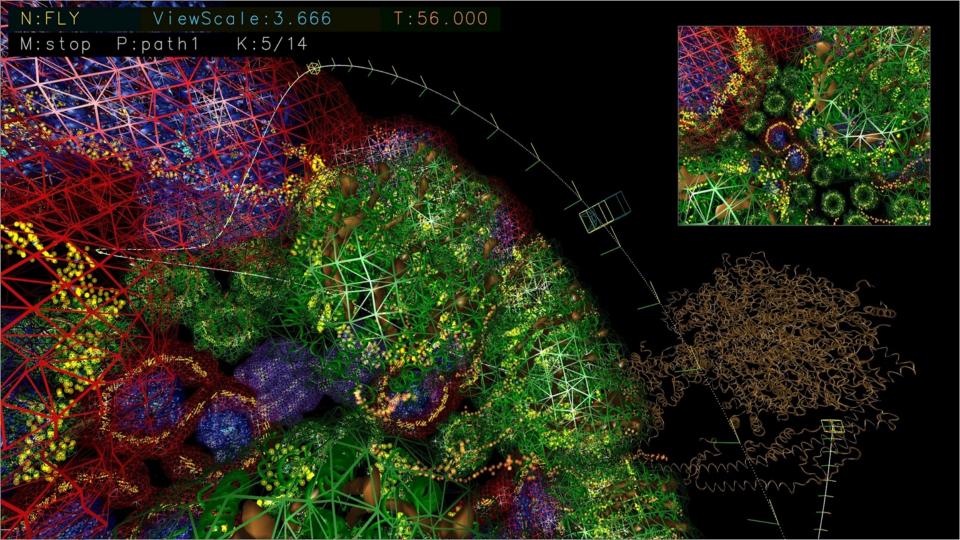




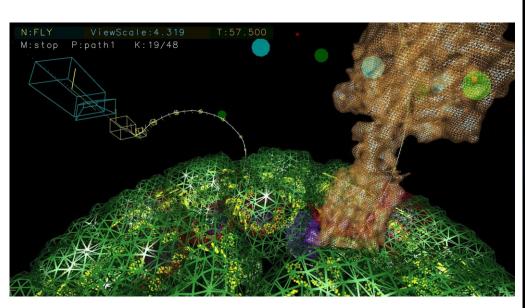


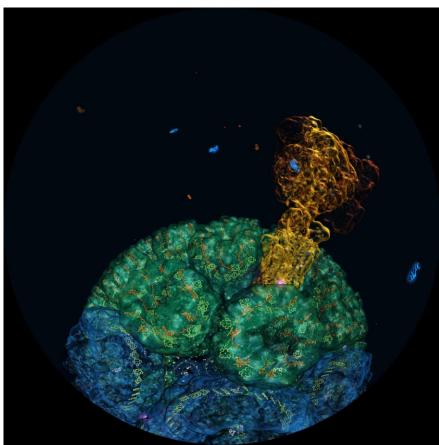
Integration of VMD with Virtual Director





#### Camera Path in Virtual Director and VMD

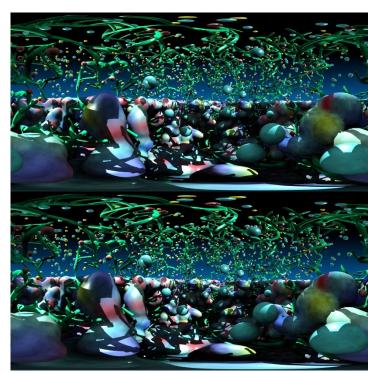




### Stereoscopic Panorama Ray Tracing



- Render 360° images and movies for VR headsets such as Oculus Rift, Google Cardboard, YouTube VR
- Ray trace omnidirectional stereo spheremaps or cubemaps for very high-frame-rate reprojection and display via OpenGL texturing
- Stereo requires spherical camera projections poorly suited to rasterization
- Benefits from OptiX multi-GPU rendering and load balancing, remote visualization



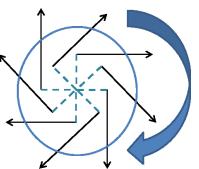
### Omnidirectional Stereoscopic Ray Tracing

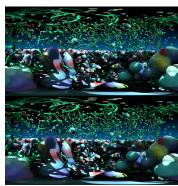
- Ray trace 360° images and movies for Desk and VR HMDs: Oculus, Vive, Cardboard
- Stereo spheremaps or cubemaps allow very highframe-rate interactive OpenGL display
- AO lighting, depth of field, shadows, transparency, curved geometry, ...
- Summit 6x Tesla V100 GPU nodes:
  - Render many omni-stereo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

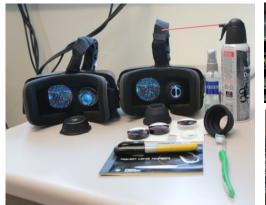
Atomic Detail Visualization of Photosynthetic Membranes with GPU-Accelerated Ray Tracing. J. E. Stone, et al. J. Parallel Computing, 55:17-27, 2016.

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.















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.

### Ray Tracing Gems

- Ch. 4, "A Planetarium Dome Master Camera"
- Ch. 27, "Interactive Ray Tracing Techniques for High-Fidelity Scientific Visualization"
- Tons of great material and code samples!

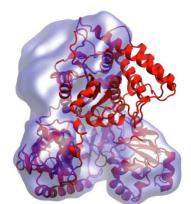


#### Molecular Dynamics Flexible Fitting (MDFF)





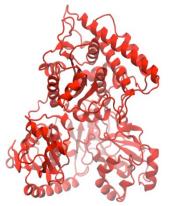
APS at Argonne



ORNL Titan



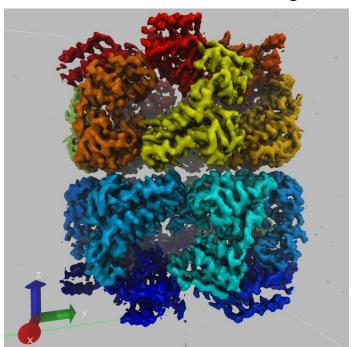
FEI microscope



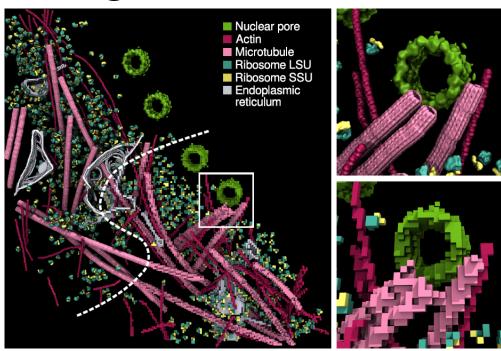


Molecular dynamics-based refinement and validation for sub-5Å cryoelectron microscopy maps. A. Singharoy, I. Teo, R. McGreevy, J. E. Stone, J. Zhao, and K. Schulten. eLife 2016;10.7554/eLife.16105

### **Density Map Segmentation**



VMD GPU-accelerated density map segmentation of GroEL



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

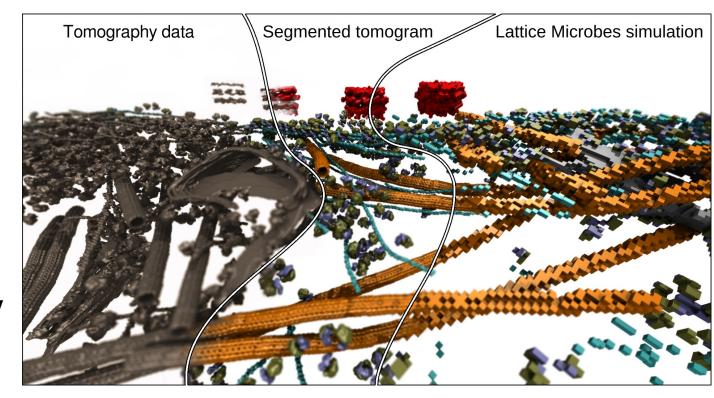




### Cryo-EM / Cryo-ET Density Map Segmentation

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...

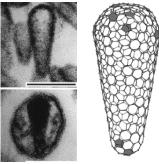




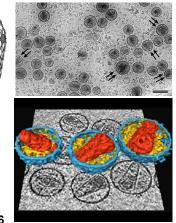


#### Structural Route to the all-atom HIV-1 Capsid

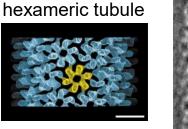
1st TEM (1999) 1st tomography (2003)



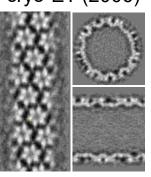
Ganser et al. *Science*, 1999 Briggs et al. *EMBO J*, 2003 Briggs et al. *Structure*, 2006



cryo-ET (2006)

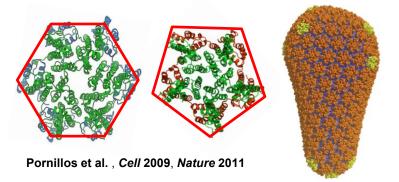


Li et al., Nature, 2000

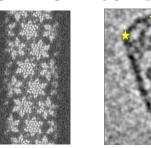


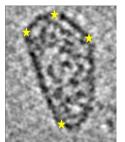
Byeon et al., Cell 2009

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



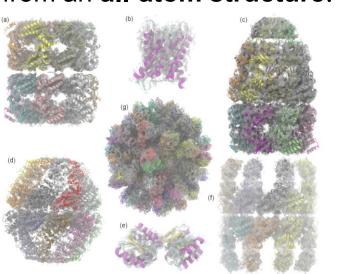


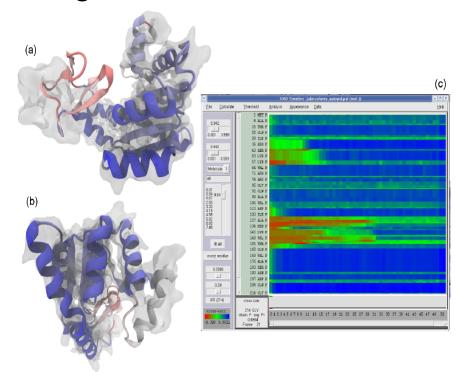


Zhao et al., *Nature* 497: 643-646 (2013)

## 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 Timeline** 

Regions with poor fit

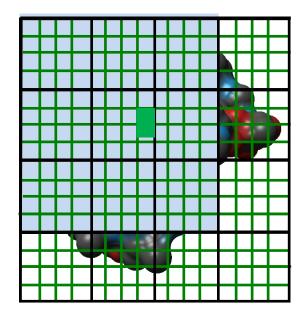
Regions with good fit

### MDFF Density Map Algorithm

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

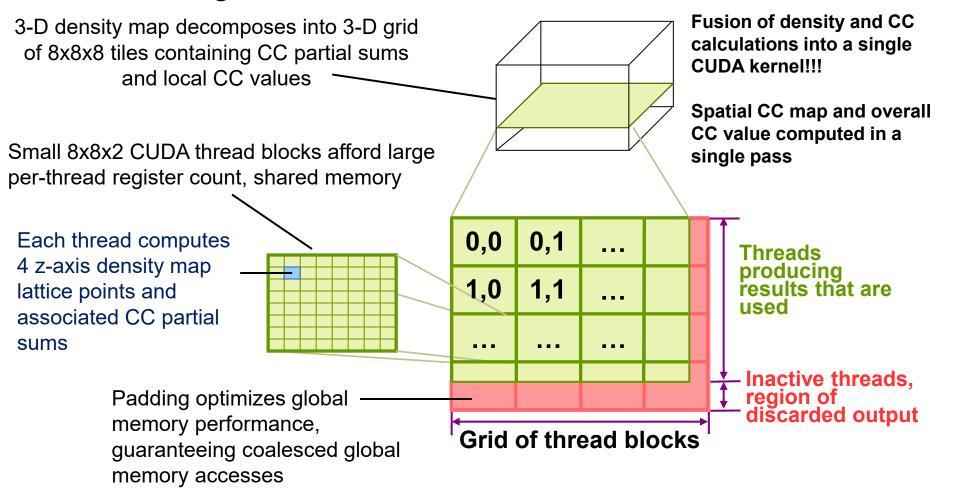
$$\rho(\vec{r}; \vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) = \sum_{i=1}^{N} e^{\frac{-|\vec{r} - \vec{r}_i|^2}{2\alpha^2}}$$

 Truncated Gaussian and spatial acceleration grid ensure linear time-complexity



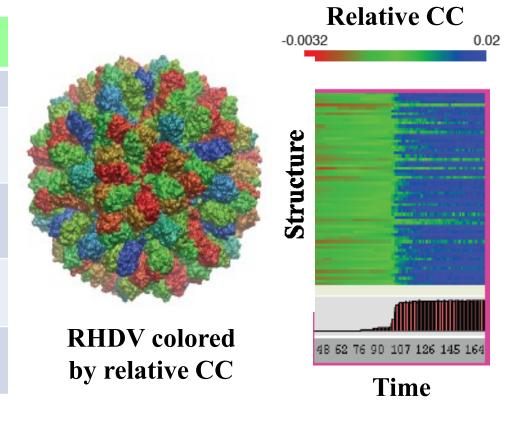
3-D density map lattice point and the neighboring spatial acceleration cells it references

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



#### Parallel MDFF Cross Correlation Analysis on Cray XK7

Rabbit Hemorrhagic Disease 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:

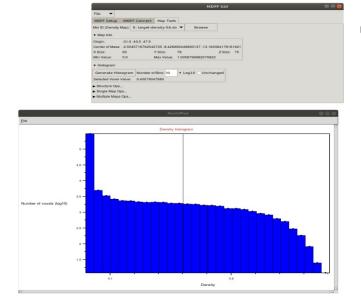
Application and Hardware platform	Runtime, Speed	lup vs. Chimera,	VMD+GPU
Chimera Xeon E5-2687W (2 socket) [1]	15.860s,	1x	
VMD-CUDA IBM Power8 + 1x Tesla K40 [2]	0.488s,	32x	0.9x
VMD-CUDA Intel Xeon E5-2687W + 1x Quadro K6000 [1,2]	0.458s,	35x	1.0x
VMD-CUDA Intel Xeon E5-2698v3 + 1x Tesla P100	0.090s,	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

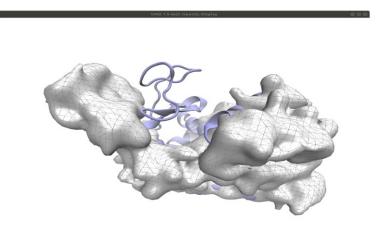
[1] GPU-Accelerated Analysis and Visualization of Large Structures Solved by Molecular Dynamics Flexible Fitting. J. E. Stone, R. McGreevy, B. Isralewitz, and K. Schulten. Faraday Discussions 169:265-283, 2014. [2] Early Experiences Porting the NAMD and VMD Molecular Simulation and Analysis Software 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.

### **New VMD MDFF Density Map Tools**

- New Map Tools tab of MDFF GUI provides wide array of density map manipulation tools including:
  - New Rigid Body Fitting
  - New Interactive Histogram
  - Trim, Crop, Clamp, Smooth...
  - Easy Masking routine

- New Density Segmentation
- Add, subtract, multiply maps
- Cross correlation and potential calculations for MDFF





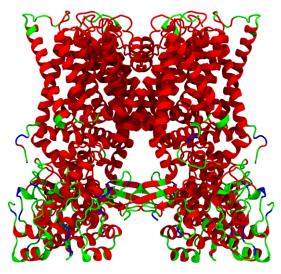




### **New Density Map Tools - Masking**

Easily select and mask density map regions with VMD selection language

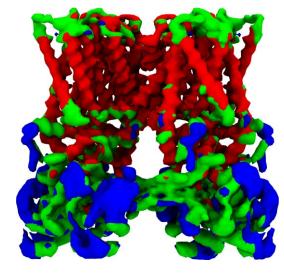
TRPV1 structure (3J5P) and cryo-EM density (emd-5778) colored by local resolution obtained by ResMap



High Res (~3 Å)

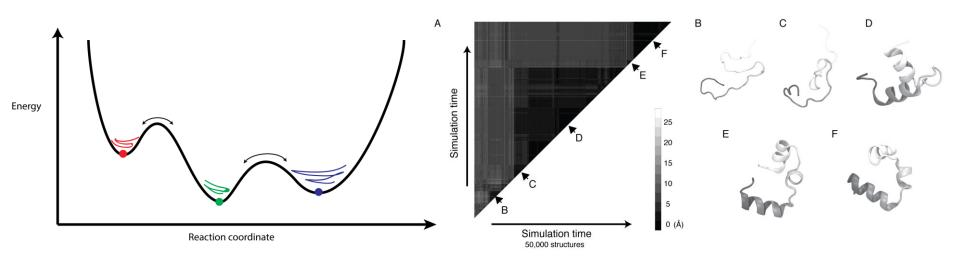
Med Res (~4 Å)

Low Res (~5 Å)



A. Kucukelbir, F.J. Sigworth, H.D. Tagare, Quantifying the Local Resolution of Cryo-EM Density Maps, Nature Methods, Volume 11, Issue 1, Pages 63-65, 2014.

## 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.** 

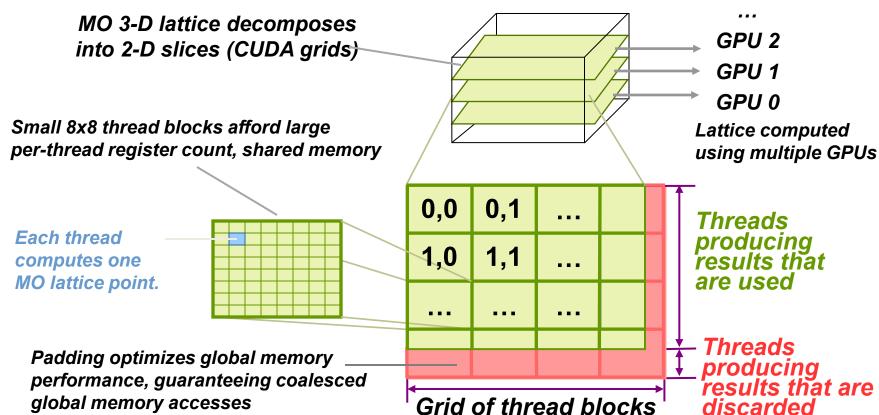
http://dx.doi.org/10.1038/nmeth.4638

**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++) {
                                                                              Loop over atoms
  int prim counter = atom basis[at];
  calc distances to atom(&atompos[at], &xdist, &ydist, &zdist, &dist2, &xdiv);
  for (contracted gto=0.0f, shell=0; shell < num shells per atom[at]; shell++) {
                                                                               Loop over shells
    int shell type = shell symmetry[shell counter];
    for (prim=0; prim < num prim per shell[shell counter]; prim++) {
                                                                              Loop over primitives:
      float exponent
                        = basis_array[prim_counter
      float contract_coeff = basis_array[prim_counter + 1];
                                                                              largest component of
       contracted gto += contract coeff * expf(-exponent*dist2);
                                                                             runtime, due to expf()
       prim counter += 2;
                                                                                 Loop over angular
    for (tmpshell=0.0f, j=0, zdp=1.0f; j<=shell type; j++, zdp*=zdist) {
      int imax = shell type - j;
                                                                                       momenta
      for (i=0, ydp=1.0f, xdp=pow(xdist, imax); i<=imax; i++, ydp*=ydist, xdp*=xdiv)
        tmpshell += wave f[ifunc++] * xdp * ydp * zdp;
                                                                               (unrolled in real code)
    value += tmpshell * contracted gto;
    shell counter++;
```

### MO GPU Parallel Decomposition







#### VMD C<sub>60</sub> MO Viz. Perf, 516x519x507 Grid:

@ .13s/frame, avoids 3.8GB/s I/O per-node

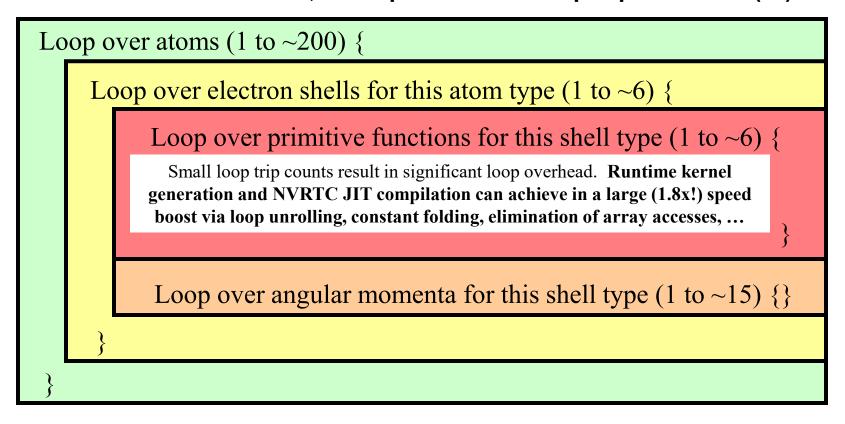
Hardware platform		Runtime,	Speedup	
IBM Power8 (ORNL 'crest') + 1x Tesla K40 [1]		3.49s,	1.0x	
Intel Xeon E5-2697Av4	+ 1x Tesla V100	0.610s,	5.7x	
Intel Xeon E5-2697Av4	+ 2x Tesla V100	0.294s,	11.8x	
Intel Xeon E5-2697Av4	+ 3x Tesla V100	0.220s,	15.9x	
IBM Power9 "Newell"	+ 1x Tesla V100	0.394s,	8.8x	7
IBM Power9 "Newell"	+ 2x Tesla V100	0.207s,	16.8x	
IBM Power9 "Newell"	+ 3x Tesla V100	0.151s,	23.1x	
IBM Power9 "Newell"	+ 4x Tesla V100	0.130s,	26.8x	

NVLink perf. boost w/ no code tuning (YET)



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

General loop-based data-dependent MO CUDA kernel

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

contracted\_gto = 1.832937 \* expf(-7.868272\*dist2); contracted\_gto += 1.405380 \* expf(-1.881289\*dist2); contracted\_gto += 0.701383 \* expf(-0.544249\*dist2);



1.8x Faster

```
for (shell=0; shell < maxshell; shell++) {
 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++) {</pre>
  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;
 float tmpshell=0;
 switch (shell type) {
  case S SHELL:
   value += const wave f[ifunc++] * contracted gto;
   break;
  case D SHELL:
   tmpshell += const wave f[ifunc++] * xdist2;
   tmpshell += const_wave_f[ifunc++] * ydist2;
   tmpshell += const wave f[ifunc++] * zdist2;
```

tmpshell += const wave f[ifunc++] \* xdist \* ydist;

# General loop-based data-dependent MO CUDA kernel

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



# 1.8x Faster

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

tmpshell = const_wave_f[ifunc++] * xdist;
tmpshell += const_wave_f[ifunc++] * ydist;
tmpshell += const_wave_f[ifunc++] * zdist;
value += tmpshell * contracted_gto;

contracted_gto = 0.187618 * expf(-0.168714*dist2);
// S_SHELL
```

value += const\_wave\_f[ifunc++] \* contracted\_gto;
contracted\_gto = 0.217969 \* expf(-0.168714\*dist2);
// P SHELL

tmpshell = const\_wave\_f[ifunc++] \* xdist;
tmpshell += const\_wave\_f[ifunc++] \* ydist;
tmpshell += const\_wave\_f[ifunc++] \* zdist;
value += tmpshell \* contracted\_gto;

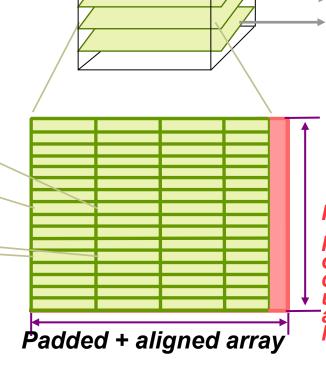
contracted\_gto = 3.858403 \* expf(-0.800000\*dist2);
// D\_SHELL
tmpshell = const\_wave\_f[ifunc++] \* xdist2;
tmpshell += const\_wave\_f[ifunc++] \* ydist2;

# MO CPU Parallel Decomposition AVX-512 Kernels on KNL

MO 3-D lattice decomposes into 2-D slices

Vectors of wavefunction amplitudes are computed using hardware SIMD instructions

Each CPU thread computes 1, 4, 8, 16
MO lattice points
per loop iteration:
C, SSE, AVX2 or
AVX-512ER



SIMD lanes producing results that are used

across many CPU

Lattice decomposed

#### **Padding:**

Thread 2

Thread 1
Thread 0

threads

Inactive SIMD lanes or region of discarded output used to guarantee aligned vector loads+stores

### 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++) {
  float exponent = basis array[prim counter
   float contract coeff = basis array[prim counter + 1];
  // contracted_qto += 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);
```





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

Intel C++ '15 autovectorization (fail):

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





# Convenient Packaging of HPC-Optimized Versions of VMD and NAMD

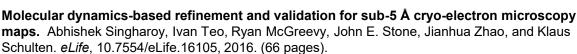




#### Making Our Research Tools Easily Accessible

- 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
- 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/



**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.

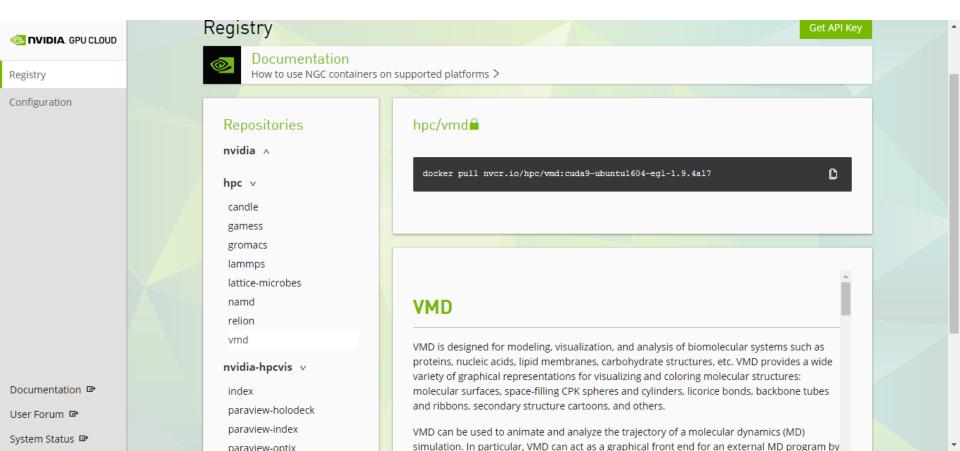


Clusters, Supercomputers



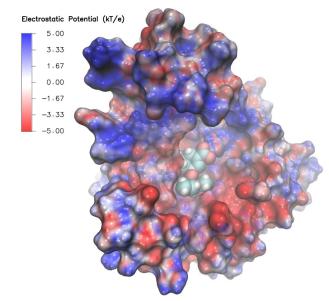


# VMD / NAMD / LM, NGC Containers



# VMD OptiX/EGL NGC Container

- https://ngc.nvidia.com/registry/
- CUDA-accelerated viz+analysis
- EGL off-screen rendering no windowing system needed
- OptiX high-fidelity GPU ray tracing engine built in
- All dependencies included
- Easy to deploy on a wide range of GPU accelerated platforms



High performance molecular visualization: In-situ and parallel rendering with EGL. J. E. Stone, P. Messmer, R. Sisneros, and K. Schulten. 2016 IEEE International Parallel and Distributed Processing Symposium Workshop (IPDPSW), pp. 1014-1023, 2016.





#### **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



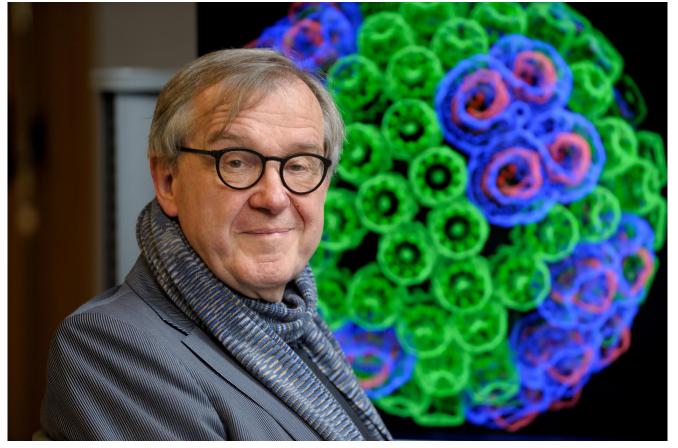
Energy efficiency: ARM+GPU

# 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

- 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. Parallel and Distributed Systems, 26(5):1405-1418, 2015. IEEE Transactions on
- Unlocking the Full Potential of the Cray XK7 Accelerator. M. D. Klein and J. E. Stone. 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.





- 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. 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.



