GPU-Accelerated Analysis of Large Biomolecular Complexes

John E. Stone

Theoretical and Computational Biophysics Group

Beckman Institute for Advanced Science and Technology

University of Illinois at Urbana-Champaign

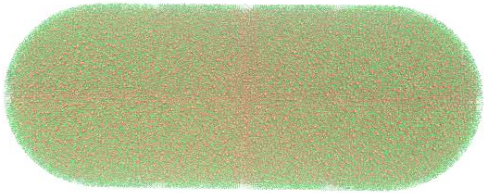
http://www.ks.uiuc.edu/

Supercomputing 2014 Exhibition

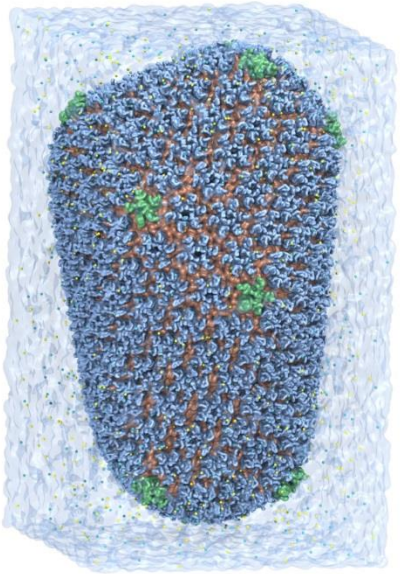
New Orleans, LA, November 18, 2014

VMD – “Visual Molecular Dynamics”

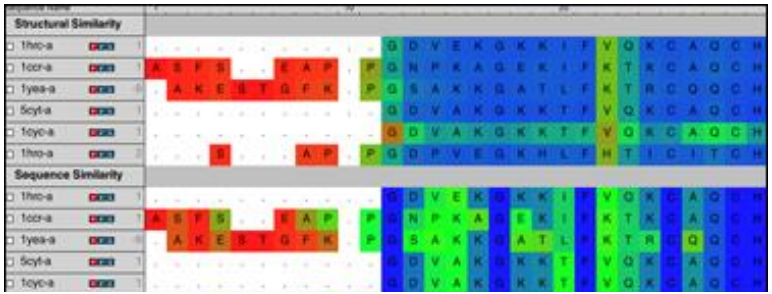
- Visualization and analysis of:
  - molecular dynamics simulations
  - particle systems and whole cells
  - cryoEM densities, volumetric data
  - quantum chemistry calculations
  - sequence information
- User extensible w/ scripting and plugins
- http://www.ks.uiuc.edu/Research/vmd/



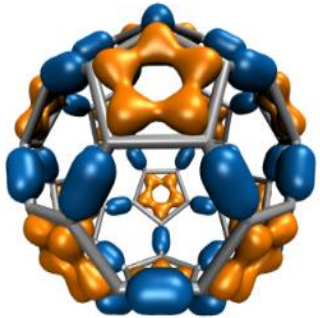
Whole Cell Simulation



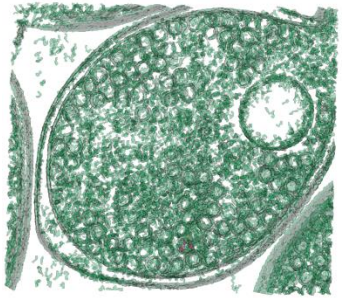
MD Simulations



Sequence Data



Quantum Chemistry

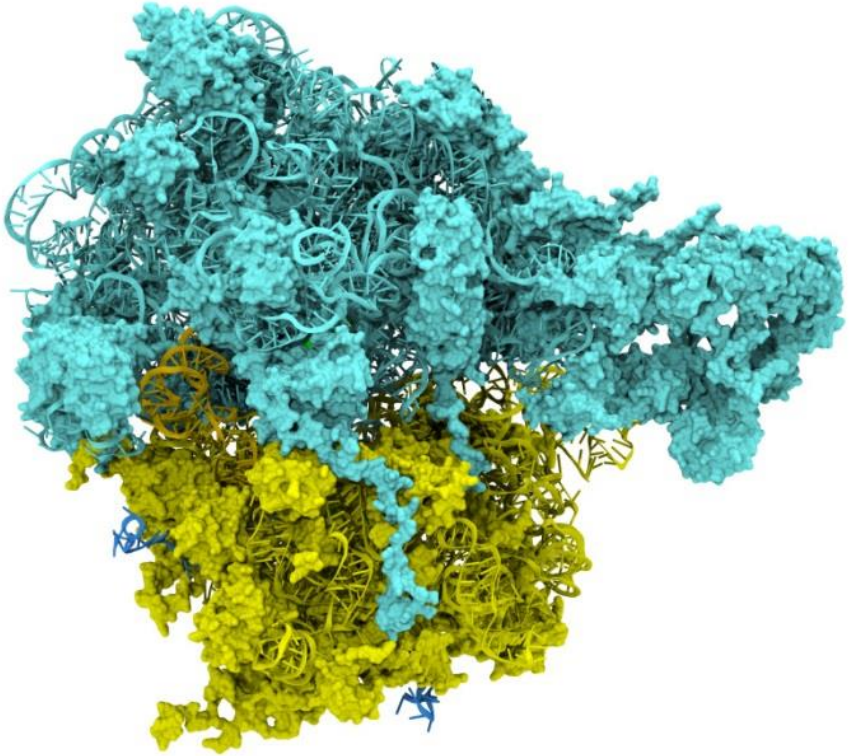


CryoEM, Cellular Tomography

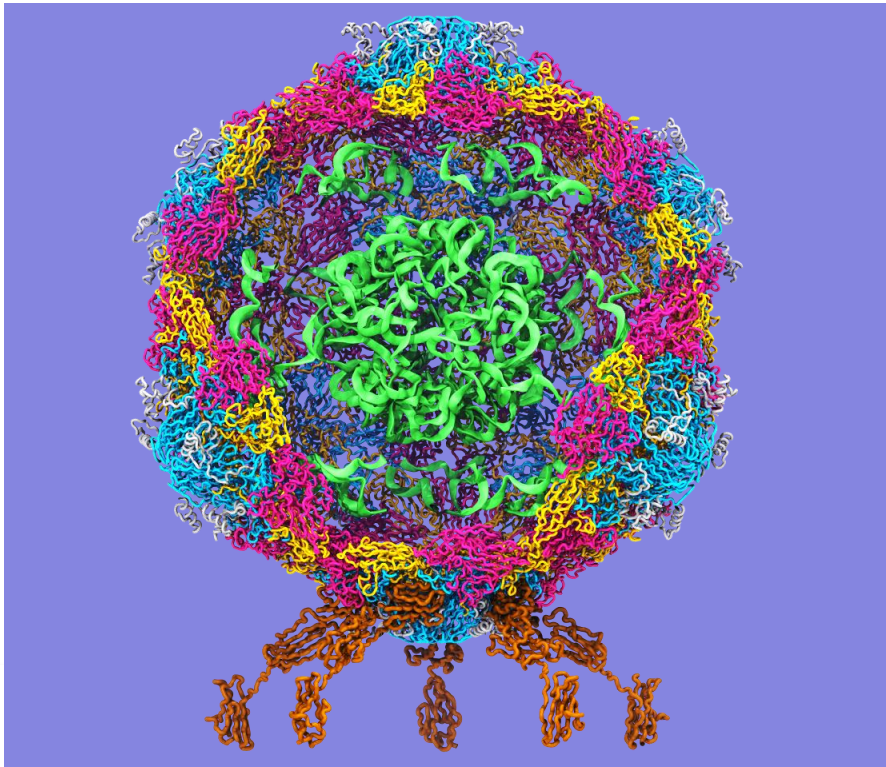
Goal: A Computational Microscope

Study the molecular machines in living cells

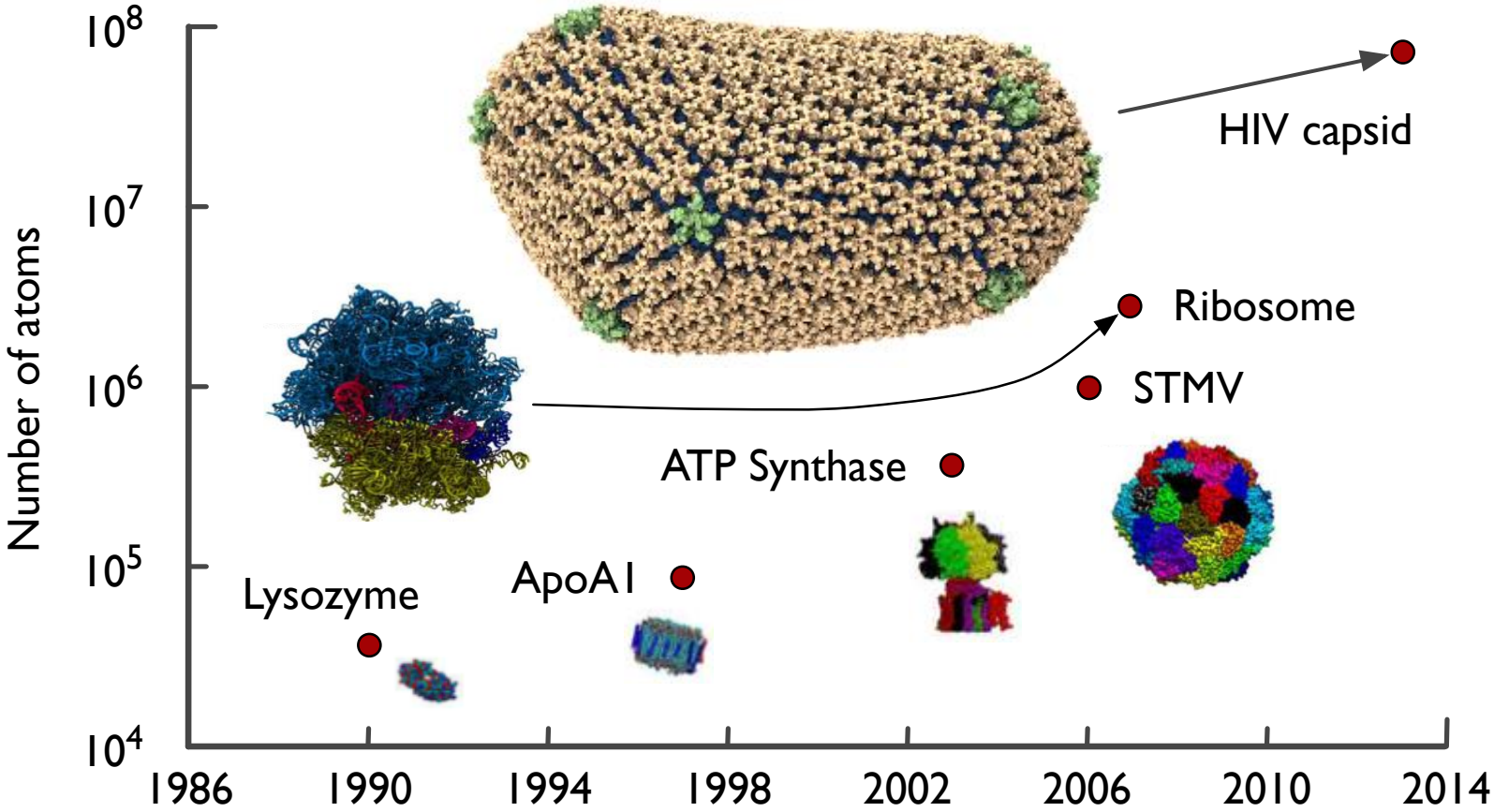
Ribosome: target for antibiotics



Poliovirus



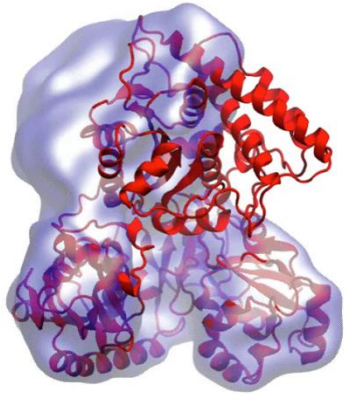
NAMD and VMD Use GPUs & Petascale Computing to Meet Computational Biology's Insatiable Demand for Processing Power



Molecular Dynamics Flexible Fitting (MDFF)

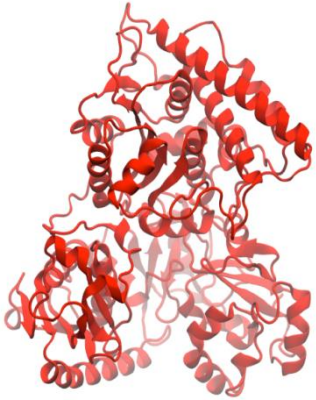
X-ray crystallography → MDFF ← Electron microscopy

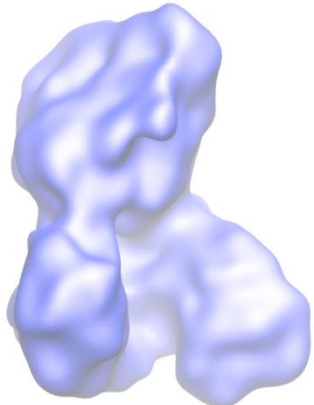
APS at Argonne



ORNL Titan

FEI microscope





Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics.
L. Trabuco, E. Villa, K. Mitra, J. Frank, and K. Schulten. Structure, 16:673-683, 2008.

Molecular Dynamics Flexible Fitting - Theory

Two terms are added to the MD potential

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

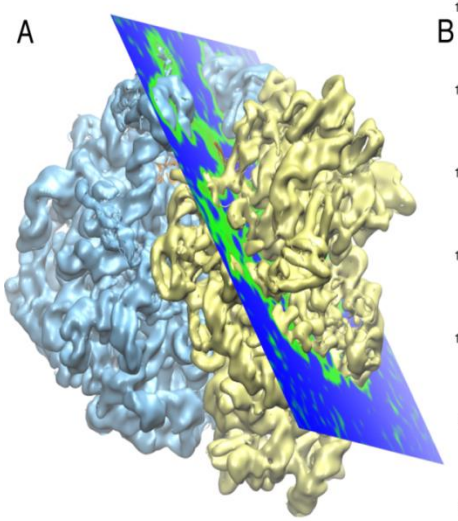
An external potential derived from the EM map is defined on a grid as

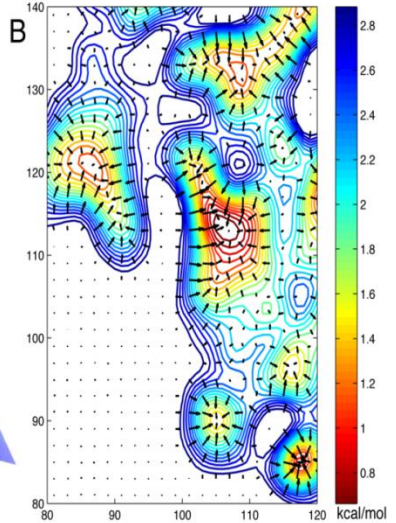
$$U_{EM}(\mathbf{R}) = \sum_j w_j V_{EM}(\mathbf{r}_j)$$

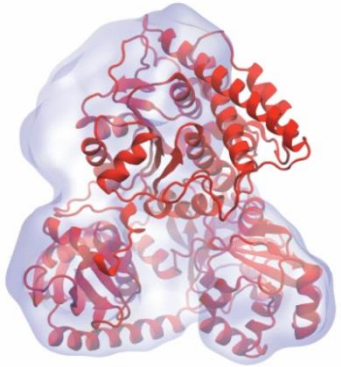
$$V_{EM}(\mathbf{r}) = \begin{cases} \xi \left(1 - \frac{\Phi(\mathbf{r}) - \Phi_{thr}}{\Phi_{max} - \Phi_{thr}}\right) & \text{if } \Phi(\mathbf{r}) \geq \Phi_{thr}, \\ \xi & \text{if } \Phi(\mathbf{r}) < \Phi_{thr}. \end{cases}$$

A mass-weighted force is then applied to each atom

$$\mathbf{f}_i^{EM} = -\nabla U_{EM}(\mathbf{R}) = -w_i \partial V_{EM}(\mathbf{r}_i)/\partial r_i$$

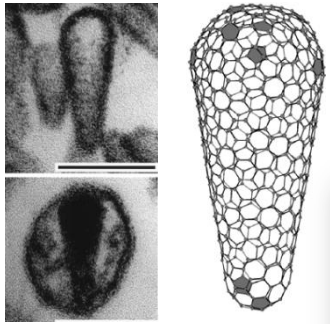






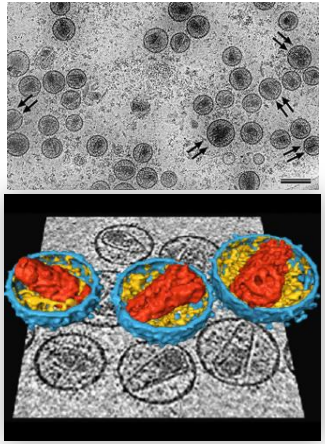
Structural Route to the HIV-1 Capsid

1st TEM (1999)

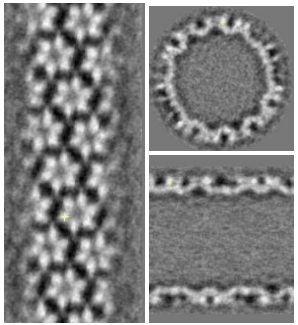


Ganser et al. *Science*, 1999
Briggs et al. *EMBO J*, 2003
Briggs et al. *Structure*, 2006

1st tomography (2003)

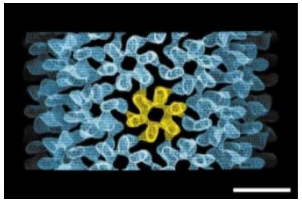


cryo-ET (2006)



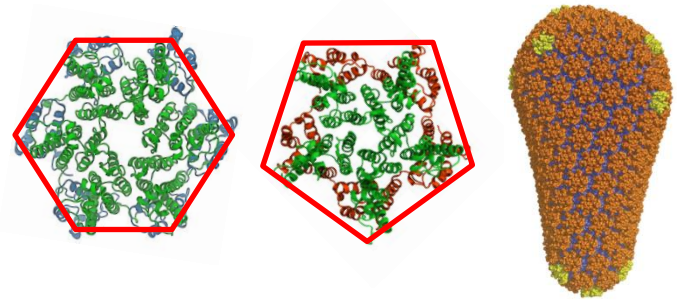
Byeon et al., *Cell* 2009

hexameric tubules



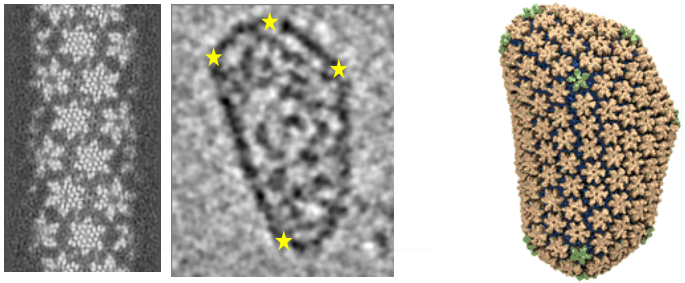
Li et al., *Nature*, 2000

Crystal structures of separated hexamer and pentamer



Pornillos et al. , *Cell* 2009, *Nature* 2011

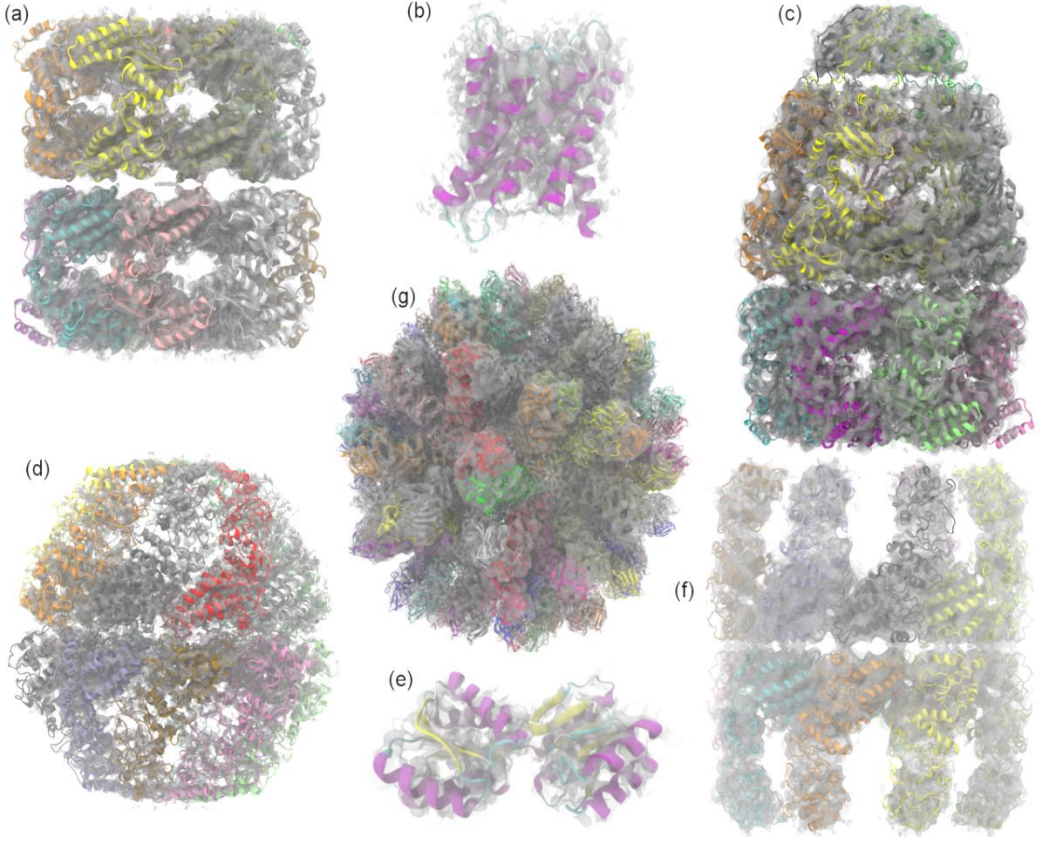
High res. EM of hexameric tubules, tomography of capsids, **all-atom model of capsid by MDFF w/ NAMD & VMD, NSF/NCSA Blue Waters petascale computer at U. Illinois**



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

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

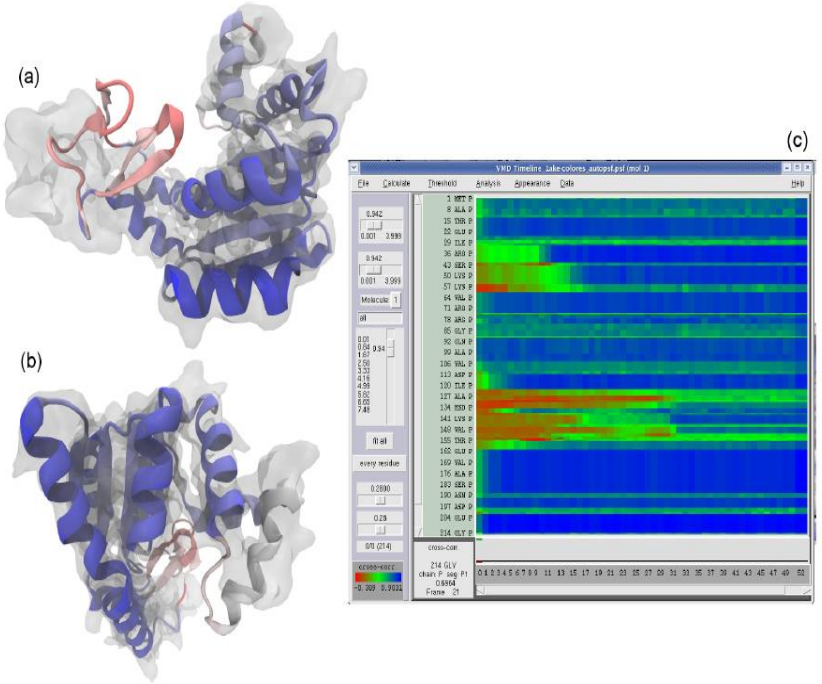
Compute Pearson correlation to evaluate the fit of a reference cryo-EM density map with a **simulated density map** produced from an **all-atom structure**.



GPUs Can Reduce MDFF Trajectory Analysis Runtimes from Hours to Minutes

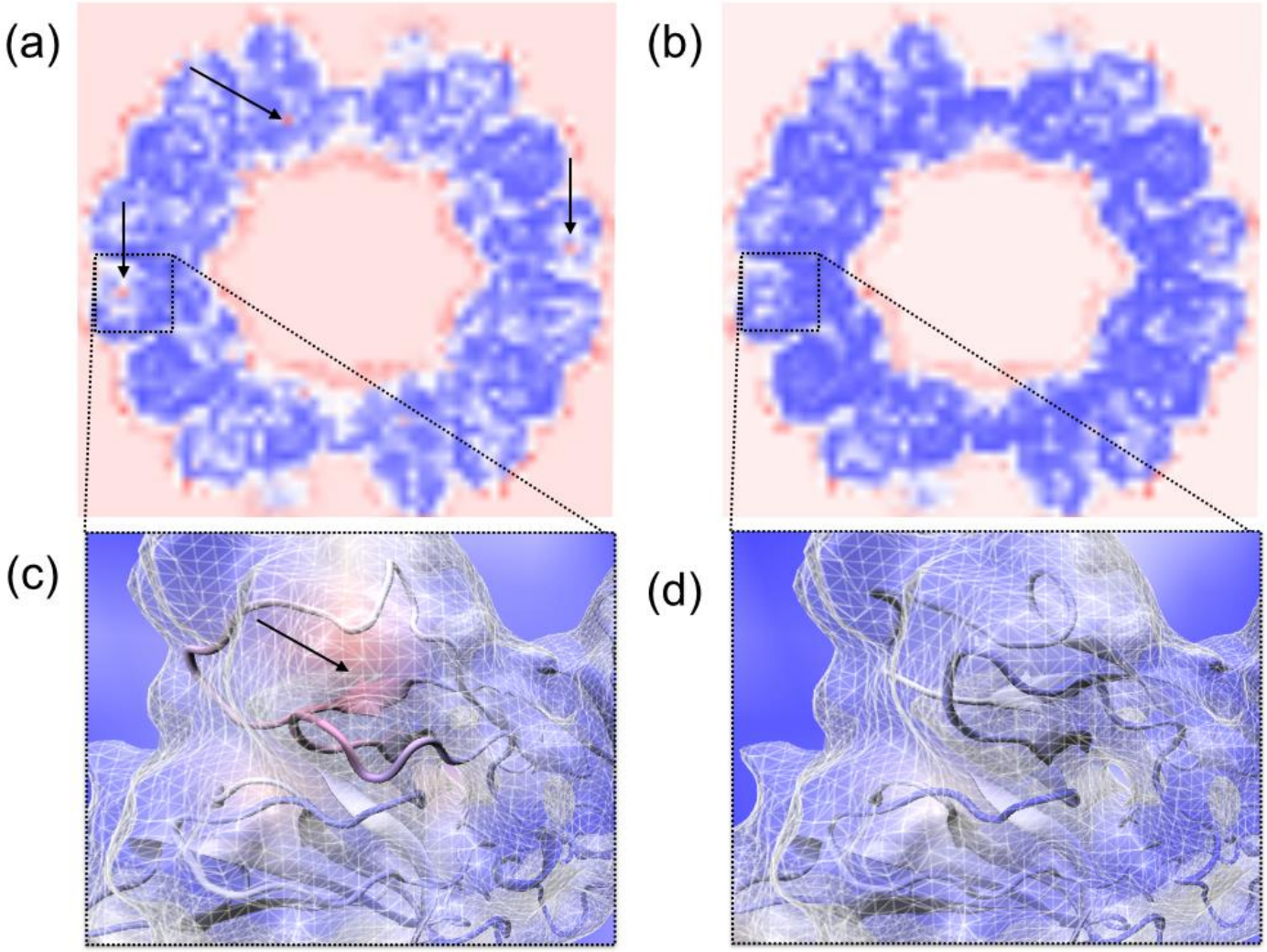
GPUs enable laptops and desktop workstations to handle tasks that would have previously required a cluster, or a *very long wait*…

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



GPU-accelerated MDFF Cross Correlation Timeline

Regions with poor fit **Regions with good fit**


(a)
(b)
(c)
(d)

Single-Pass MDFF GPU Cross-Correlation

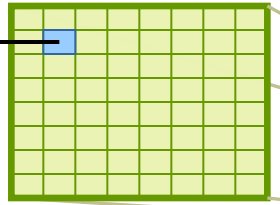
3-D density map decomposes into 3-D grid of 8x8x8 tiles containing CC partial sums and local CC values

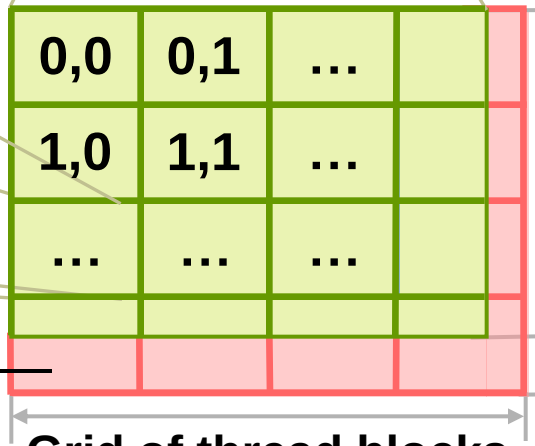
Fusion of density and CC calculations into a single CUDA kernel!!!

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

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

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





Threads producing results that are used

Inactive threads, region of discarded output

Grid of thread blocks

Padding optimizes global memory performance, guaranteeing coalesced global memory accesses

VMD GPU Cross Correlation Performance

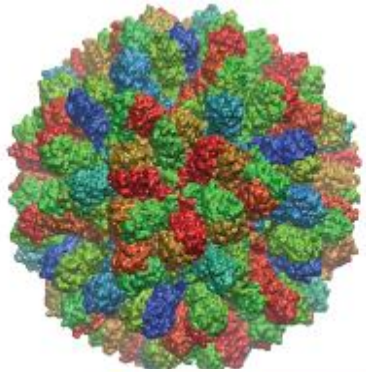
	RHDV	Mm-cpn open	GroEL	Aquaporin
Resolution (Å)	**6.5**	**8**	**4**	**3**
Atoms	**702K**	**61K**	**54K**	**1.6K**
VMD-CUDA Quadro K6000	**0.458s 34.6x**	**0.06s 25.7x**	**0.034s 36.8x**	**0.007s 55.7x**
VMD-CPU-SSE 32-threads, 2x Xeon E5-2687W	**0.779s 20.3x**	**0.085s 18.1x**	**0.159s 7.9x**	**0.033s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

GPU-Accelerated Analysis and Visualization of Large Structures Solved by Molecular Dynamics Flexible Fitting. J. E. Stone, R. McGreevy, B. Isralewitz, and K. Schulten. Faraday Discussions 169:265-283, 2014.

VMD RHDV Cross Correlation Timeline on Cray XK7

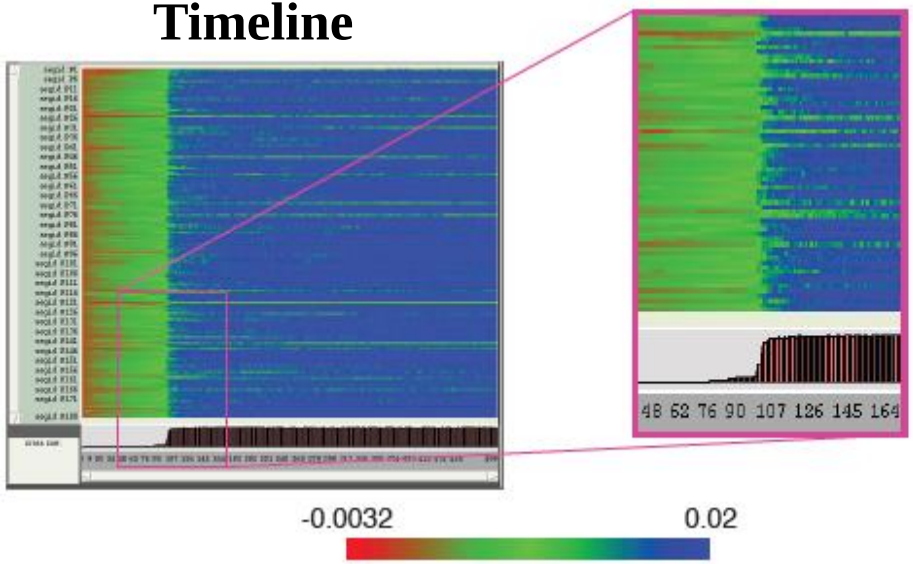
	RHDV
Atoms	702K
Traj. Frames	10,000
Component Selections	720
Single-node XK7 (projected)	336 hours (14 days)
128-node XK7	3.2 hours 105x speedup
2048-node XK7	19.5 minutes 1035x speedup

Calculation would take **5 years** using original serial CC calculation on a workstation!

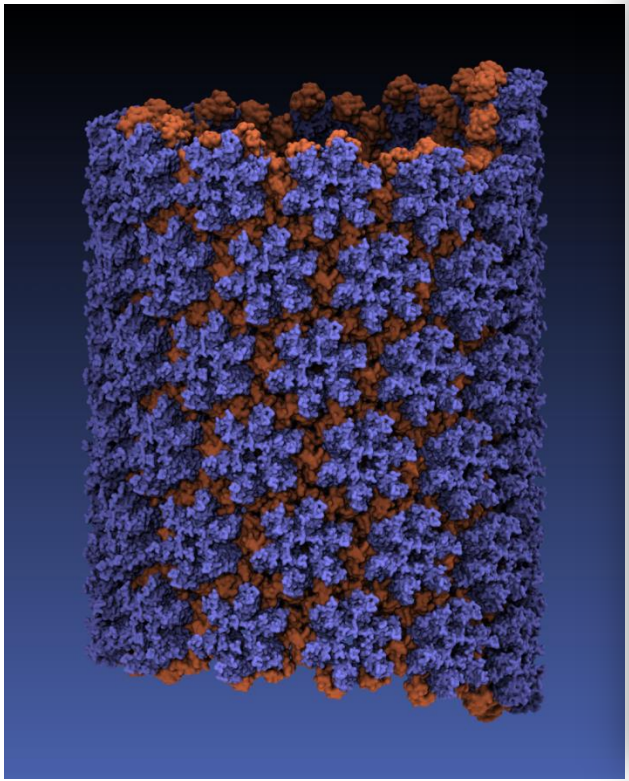


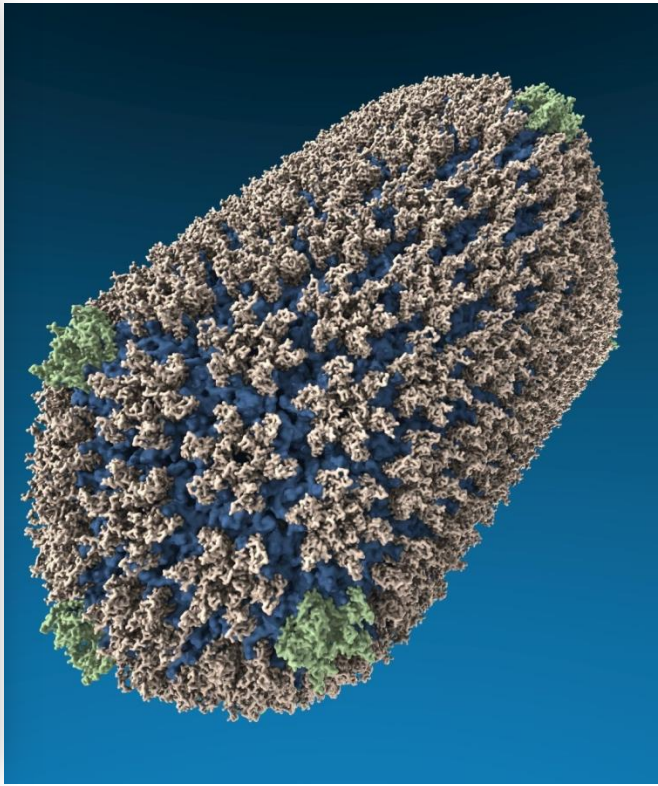
RHDV

Group-relative CC Timeline



VMD GPU-Accelerated Ray Tracing





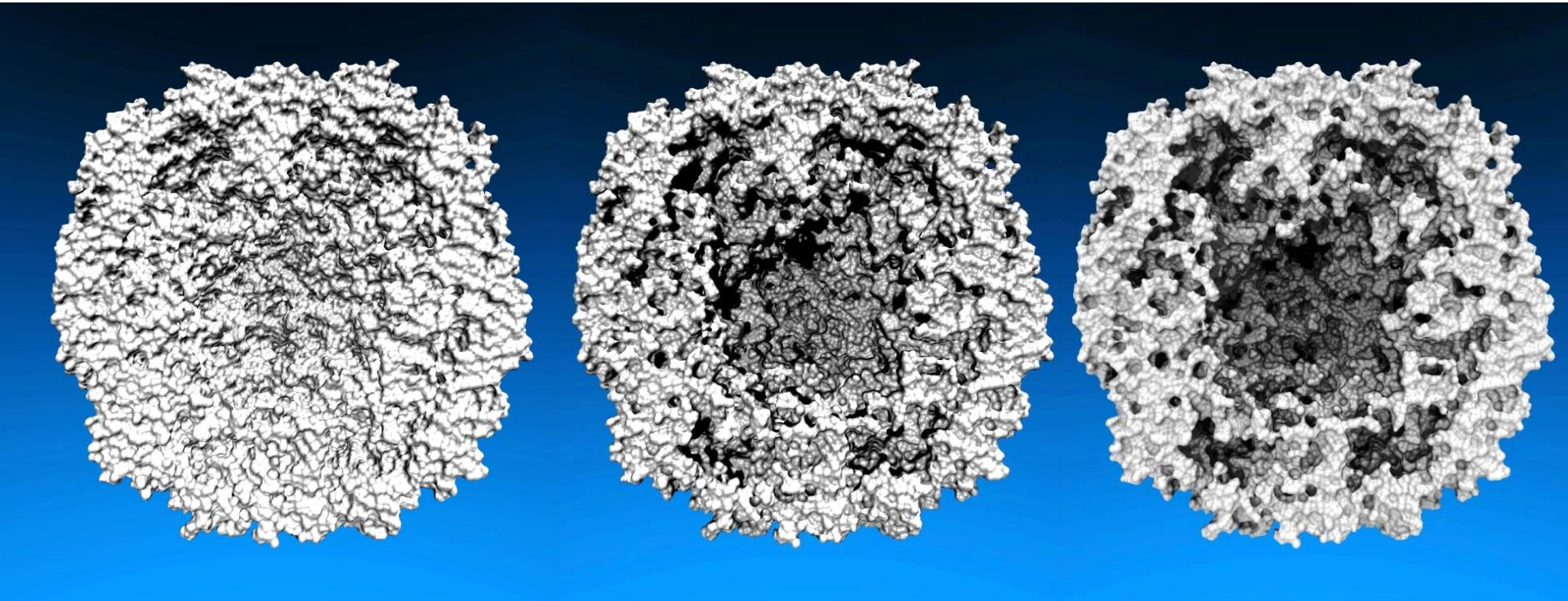
All-atom HIV capsid simulations w/ up to 64M atoms on Blue Waters

Lighting Comparison

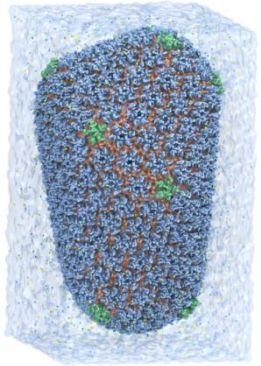
Two lights, no shadows

Two lights, hard shadows, 1 shadow ray per light

Ambient occlusion + two lights, 144 AO rays/hit



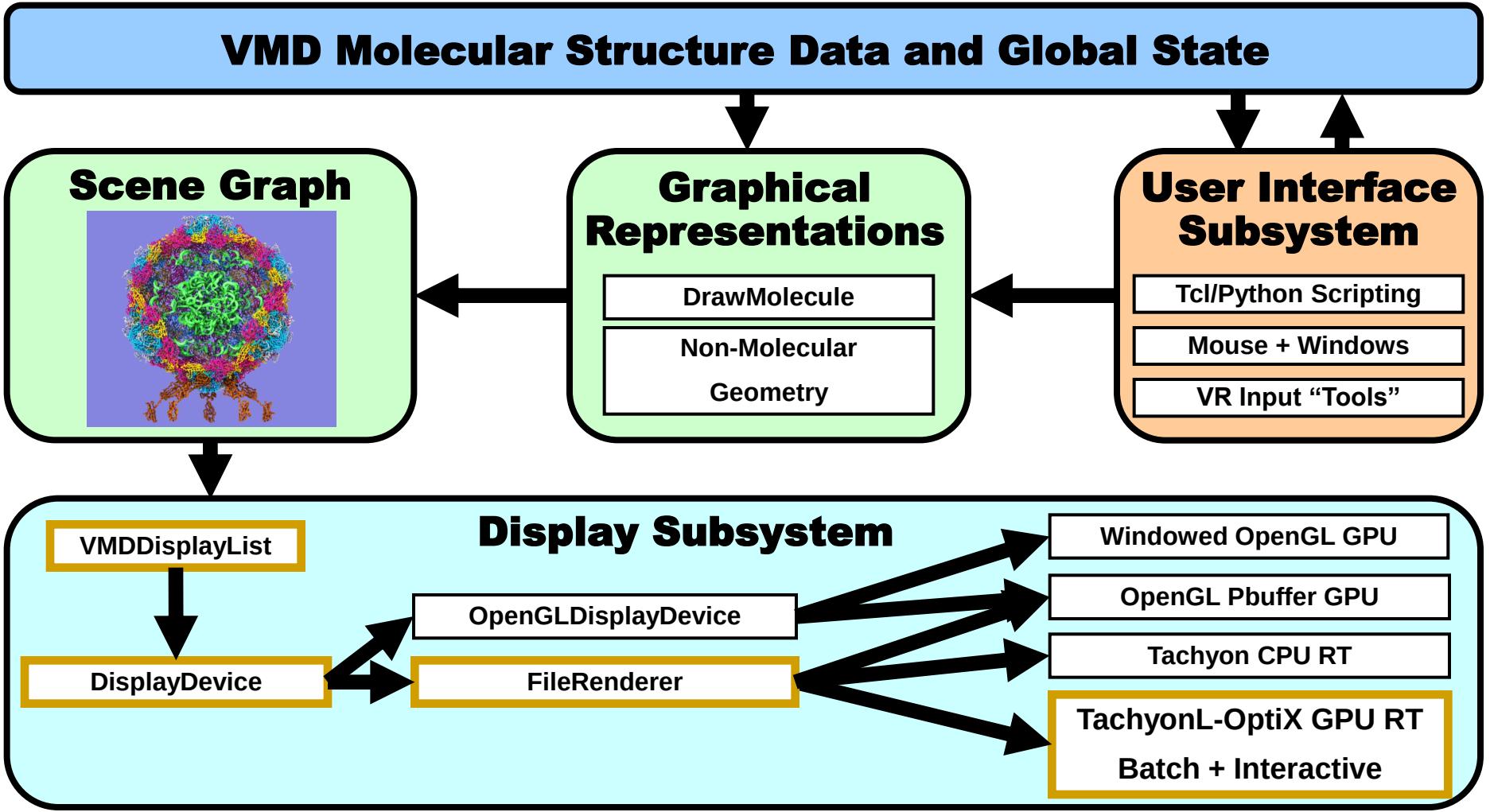
HIV-1 Parallel HD Movie Rendering on Blue Waters Cray XE6/XK7



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

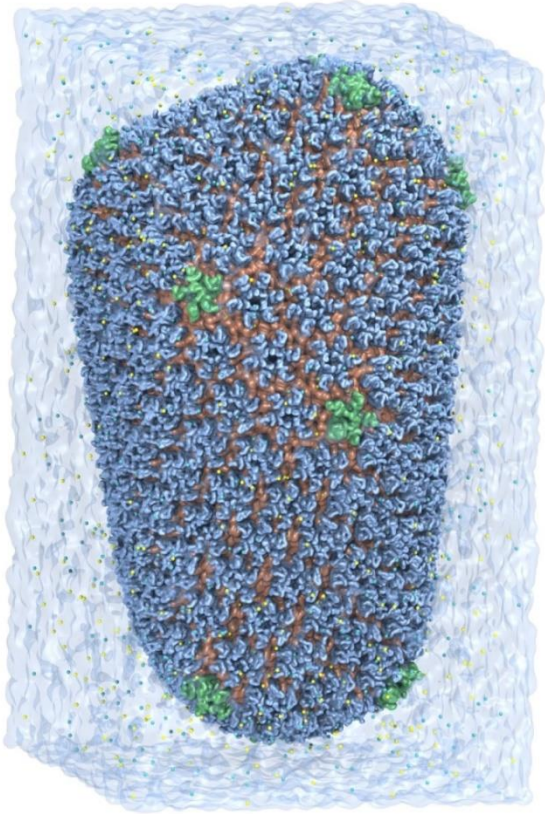
Node Type and Count	Script Load Time	State Load Time	Geometry + Ray Tracing	Total Time
256 XE6 CPUs	7 s	160 s	**1,374 s**	**1,541 s**
512 XE6 CPUs	13 s	211 s	808 s	1,032 s
64 XK7 Tesla K20X GPUs	2 s	38 s	655 s	695 s
128 XK7 Tesla K20X GPUs	4 s	74 s	331 s	410 s
256 XK7 Tesla K20X GPUs	7 s	110 s	**171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

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

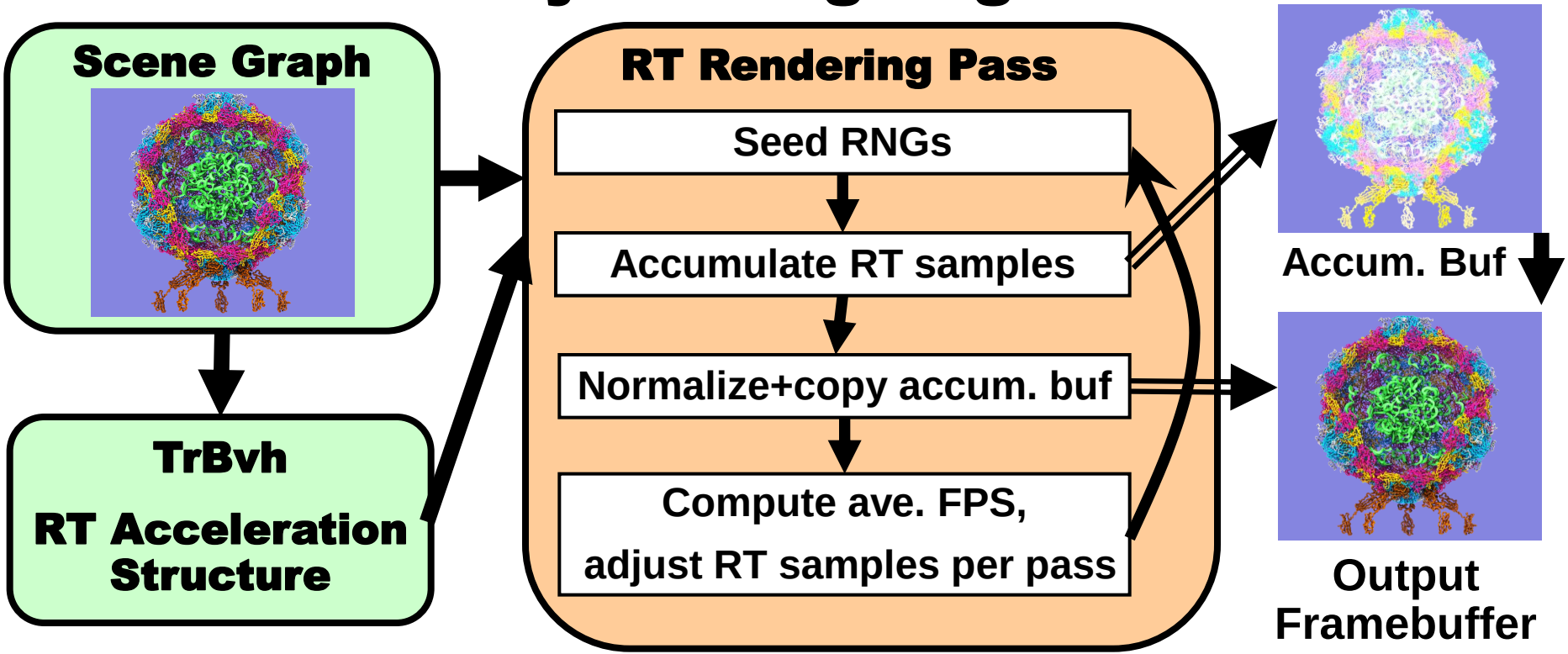

VMD Molecular Structure Data and Global State
Scene Graph
Graphical Representations
DrawMolecule
Non-Molecular Geometry
User Interface Subsystem
Tcl/Python Scripting
Mouse + Windows
VR Input “Tools”
Display Subsystem
VMDDisplayList
DisplayDevice
OpenGLDisplayDevice
FileRenderer
Windowed OpenGL GPU
OpenGL Pbuffer GPU
Tachyon CPU RT
TachyonL-OptiX GPU RT
Batch + Interactive

VMD 1.9.2 Interactive GPU Ray Tracing

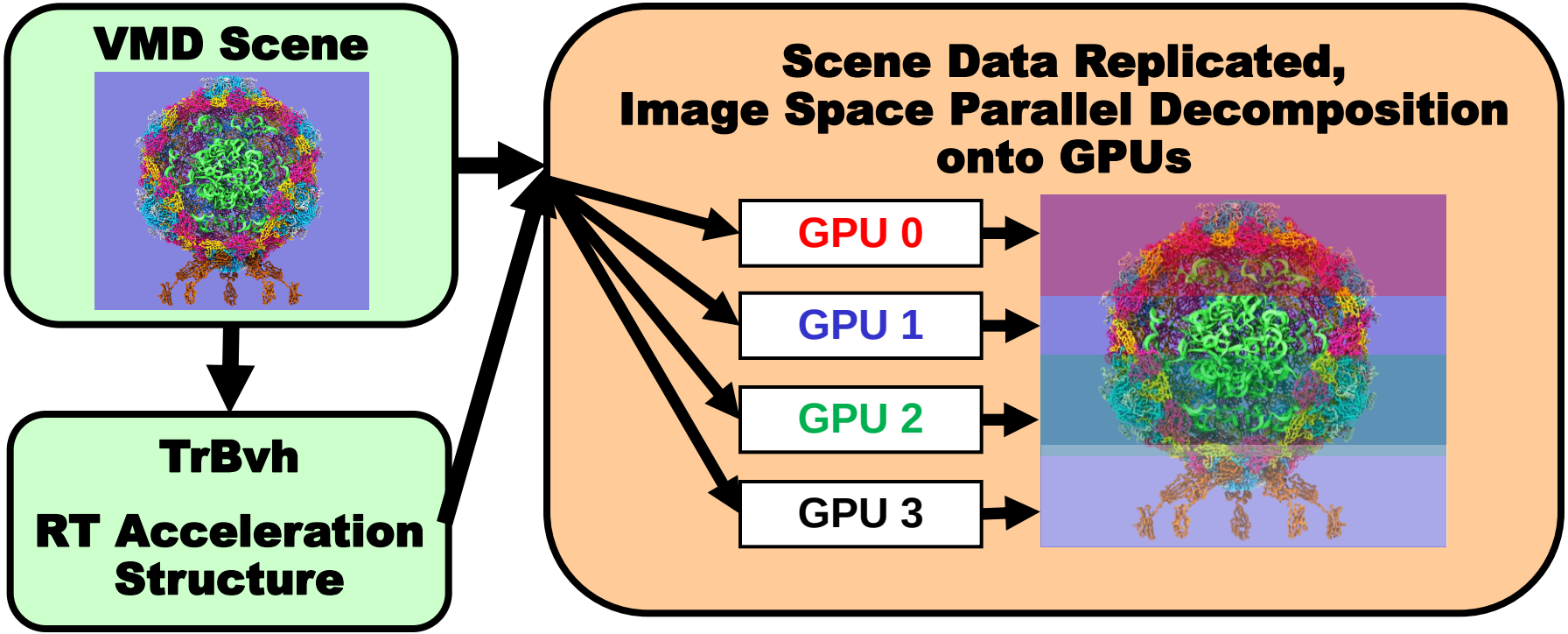
- Ray tracing heavily used for VMD publication-quality images/movies
- High quality lighting, shadows, transparency, depth-of-field focal blur, etc.
- VMD now provides –***interactive***– ray tracing on laptops, desktops, and ***remote*** visual supercomputers



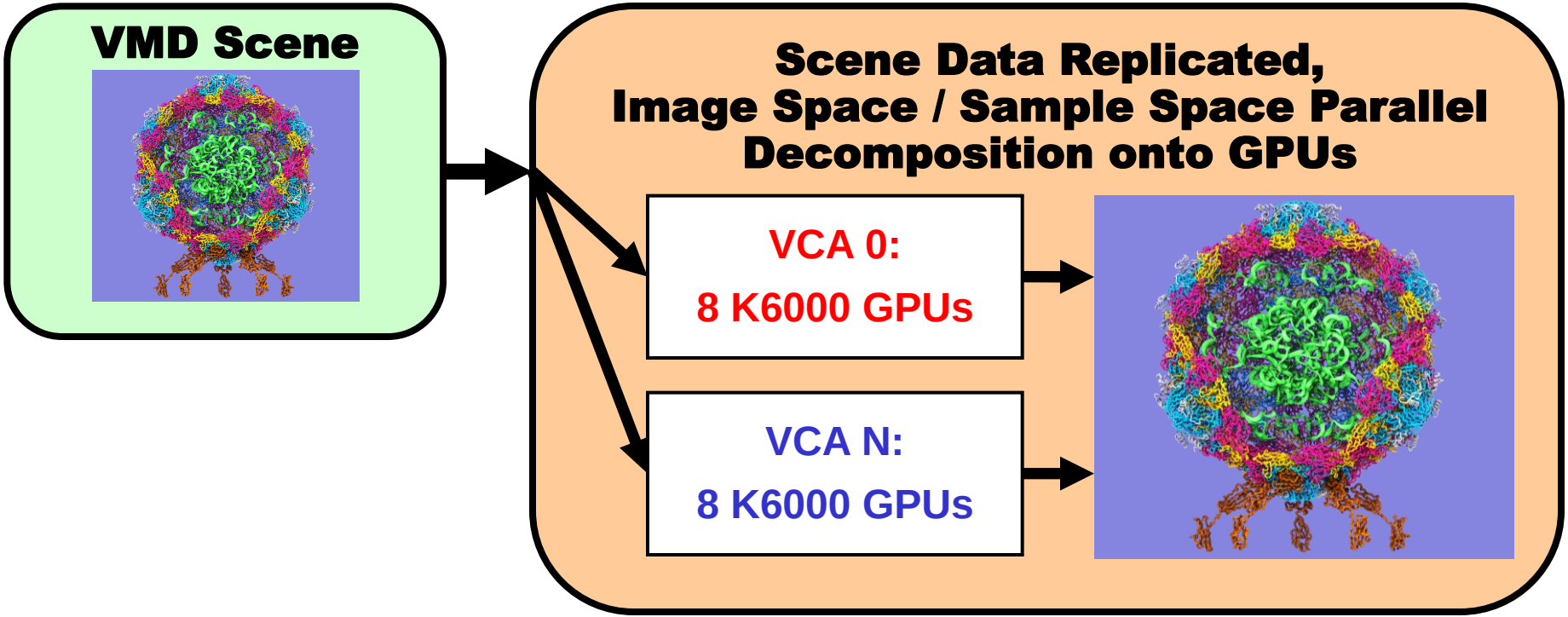
VMD TachyonL-OptiX Interactive Ray Tracing Engine



VMD TachyonL-OptiX: Multi-GPU on a Desktop or Single Node



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



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

See the live demos!

Acknowledgements

- Theoretical and Computational Biophysics Group, University of Illinois at Urbana-Champaign
- NVIDIA CUDA Center of Excellence, University of Illinois at Urbana-Champaign
- NVIDIA OptiX Team
- Funding:
  - NIH support: 9P41GM104601, 5R01GM098243-02
  - NSF PRAC “The Computational Microscope”, OCI-0832673 and ACI-1440026, and Blue Waters OCI 07-25070 and ACI-1238993
  - DOE INCITE DE-AC05-00OR22725


NIH BTRC for Macromolecular Modeling and Bioinformatics
1990-2017
Beckman Institute
University of Illinois at
Urbana-Champaign

GPU Computing Publications

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

- **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
  SC'14 Visualization and Data Analytics Showcase, 2014. **(In press)**
  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, 2014. **(In press)**
- **Unlocking the Full Potential of the Cray XK7 Accelerator.** M. D. Klein and J. E. Stone.
  Cray Users Group, Lugano Switzerland, 2014. **(In press)**
- **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.