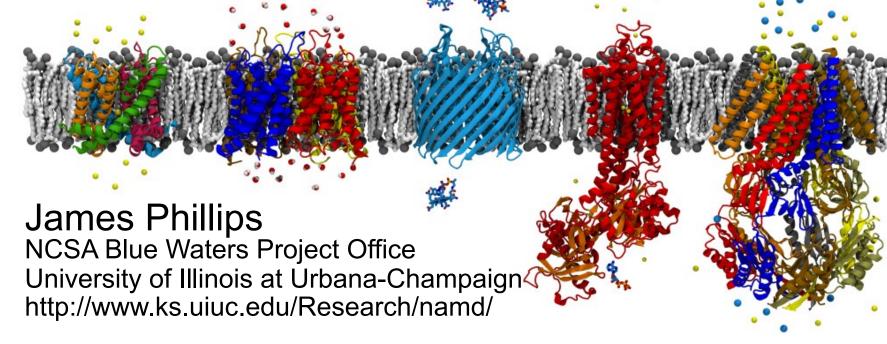
Experiences with Charm++ and NAMD on the Summit POWER9/Volta Supercomputer





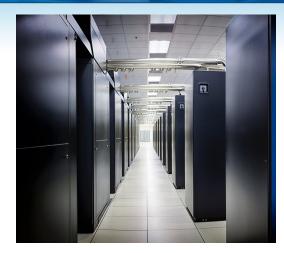






The Blue Waters Project

- Comprehensive development, deployment and service phases with co-design etc.
- The Blue Waters system is a top ranked system in all aspects of its capabilities.
- Diverse Science teams are able to make excellent use of those capabilities due to the system's <u>flexibility</u> and emphasis on sustained performance.
 - 45% larger than any system Cray has ever built
 - 22,640 CPU-only nodes, 4,224 GPU-accelerated nodes
 - Peak performance and delivered cycles are approximately the same as the aggregate of all the NSF XSEDE resources.
 - Ranks in the top systems in the world in peak performance despite being over five years old
 - Largest memory capacity (1.66 PetaBytes) of any HPC system in the world! One of the fastest file systems (>1 TB/s) in the world!
 - Largest certified nearline tape system (>250 PB) in the world
 - Fastest external network capability (>420 Gb/s) of any open science site.





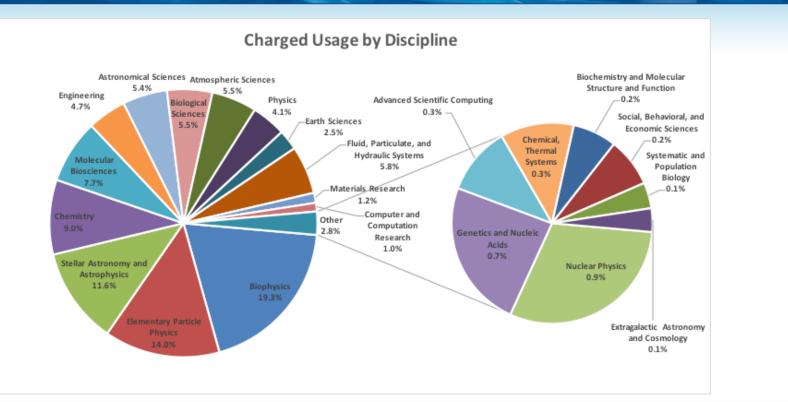


Towards a Leadership-Class Computing Facility - Phase 1

- "robust, well-balanced, and forward-looking computational asset for a broad range of research topics for which advances in fundamental understanding require the most extreme computational and **data analysis** capabilities"
- "at least two- to three-fold time-to-solution performance improvement over the current state of the art, the University of Illinois at Urbana-Champaign's (UIUC) Blue Waters system, for a broad range of existing and emerging computational and data intensive applications;"
- "scientific and technical evaluation of the Phase 1 system that will lead to an upgrade design of a leadership-class system, called the Phase 2 system,"
- "the Phase 2 system is expected to have a ten-fold or more time-to-solution performance improvement over the Phase 1 system;"
- See https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503148

All comments are personal views only and do not represent NCSA.

BLUE WATERS



CBA

GREAT LAKES CONSORTIUM

FOR PETASCALE COMPUTATION

NESA

Data From Blue Waters 2015-2016 Annual Report

NIH Biomedical Technology Research Center for Macromolecular Modeling and Bioinformatics

Developers of the widely used computational biology software VMD and NAMD

250,000 registered VMD users 80,000 registered NAMD users

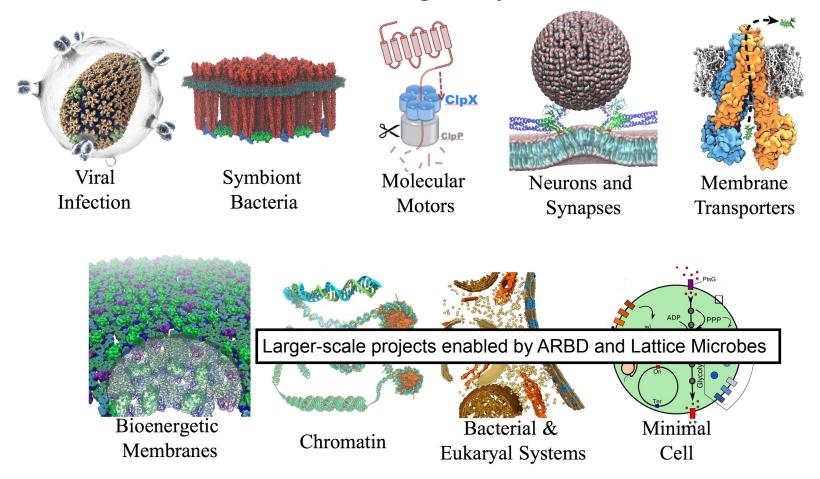
600 publications (since 1972) over 54,000 citations

4 faculty members
8 developers
1 systems
administrator
17 postdocs
46 graduate students
2 administrative staff

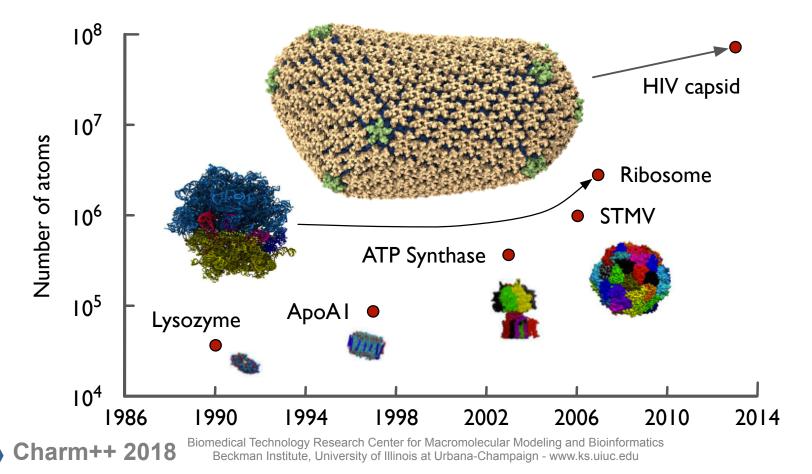
Perfect score (1<mark>0.0) on</mark> 2017-2022 NIH renewal

research projects include: virus capsids, bacteria, molecular motors, neurons and synapses, membrane transporters, bioenergetic membranes

NIH Center Driving Projects 2017-2022



Need for petascale: Simulation follows structural discovery



NAMD: Practical Supercomputing for Biomedical Research

"widest-used application" on NCSA Blue Waters, NSF-specified benchmark for successor machine

"by a very large margin the most used code" at Texas Advanced Computing Center (2nd largest)

Early adopters of workstation clusters (1993), Linux clusters (1998), and CUDA (**2007**).

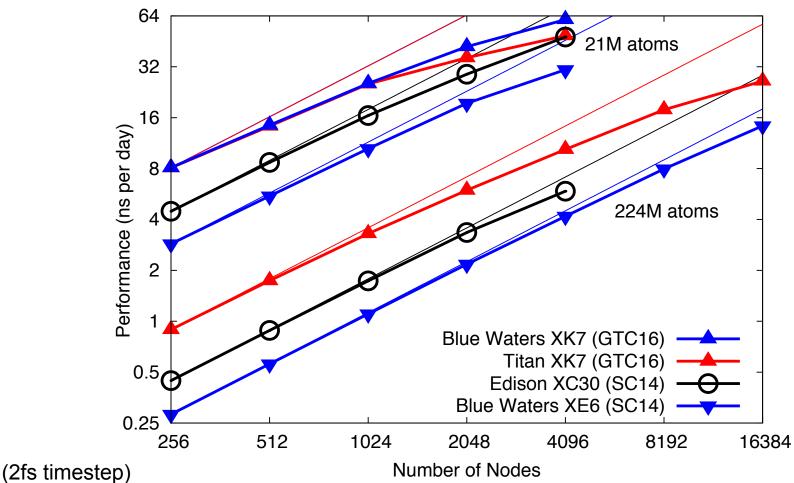
Application readiness/early science projects on

- Argonne Theta (10 PF Cray KNL, completed)
- Oak Ridge Summit (200 PF Power9/Volta, 2018)
- Argonne Aurora (200 PF Cray KNH, 2019)
- Argonne Aurora (1 EF Intel ???, 2021)



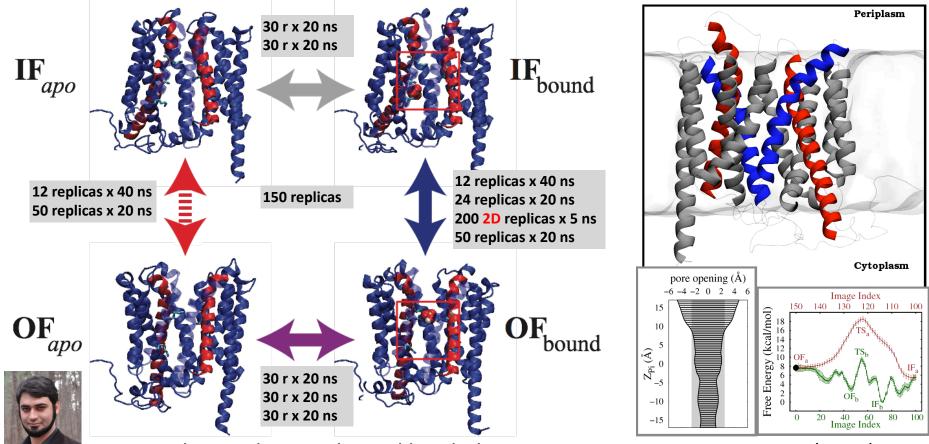
"For outstanding contributions to the development of widely used parallel software for large biomolecular systems simulation"

NAMD Runs Large Simulations Well



Multi-copy methodologies enable study of millisecond processes

Bias-exchange umbrella sampling simulations of GIpT membrane transporters



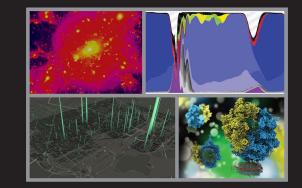
M. Moradi, G. Enkavi, and E. Tajkhorshid, Nature Communications 6, 8393 (2015)

NAMD is based on Charm++

- Parallel C++ with *data driven* objects.
- Asynchronous method invocation.
- Prioritized scheduling of messages/execution.
- Measurement-based load balancing.
- Portable messaging layer.

Complete info at charmplusplus.org and charm.cs.illinois.edu SERIES IN COMPUTATIONAL PHYSICS Steven A. Gottlieb and Rubin H. Landau, Series Editors

Parallel Science and Engineering Applications The Charm++ Approach

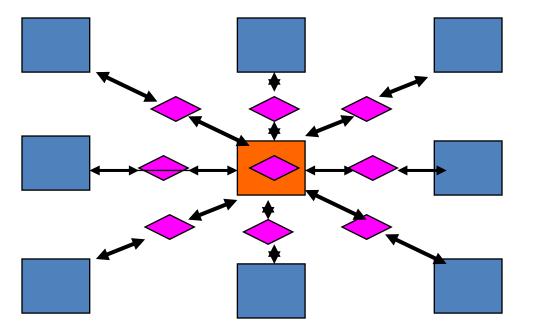


Edited by Laxmikant V. Kale Abhinav Bhatele



NAMD Hybrid Decomposition

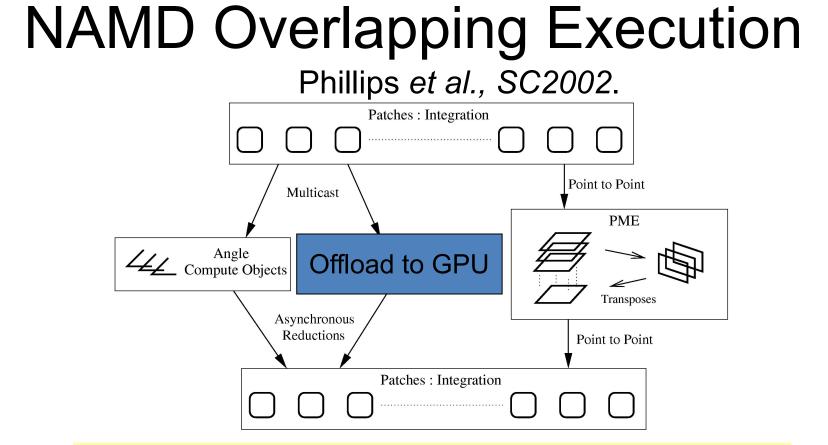
Kale et al., J. Comp. Phys. 151:283-312, 1999.



- Spatially decompose data and communication.
- Separate but related work decomposition.
- "Compute objects" facilitate iterative, measurement-based load balancing system.



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



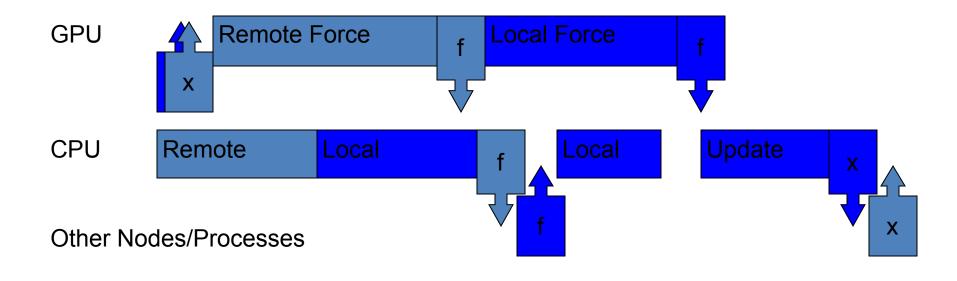
Objects are assigned to processors and queued as data arrives.



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



Overlapping GPU and CPU with Communication



One Timestep

Charm++ 2018

Phillips et al., SC2008



Streaming CPU Results to CPU

- Allows incremental results from a single grid to be processed on CPU before grid finishes on GPU
- Allows merging and prioritizing of remote and local work
- GPU side:
 - Write results to host-mapped memory (also without streaming)
 - __threadfence_system() and __syncthreads()
 - Atomic increment for next output queue location
 - Write result index to output queue
- CPU side:
 - Poll end of output queue (int array) in host memory

Non-Streaming Kernel

29.833.000 29.834.000 29.835.000 29.835.000 29.837.000 29.838.000 29.839.000 29.840.000 29.841.000 29.842.000 29.829.000 29,830,000 29.831.000 29.832.000 29.843.000 29.844.000 29.845.000 29.846.000 29.847.000 29 848 00

and the second	مار می ماند. به روی و در <mark>است</mark> اب <mark>بر از این است. از این این این این این این این این این این</mark>	
	an se a se	بىرىرىرى بەركى بەركى يەركى تەك ەر ب ەركى بەركى يەركى بەركى
		na inina di anti anti anti anti anti anti anti ant
		······································

NIH Charm++ 2018

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



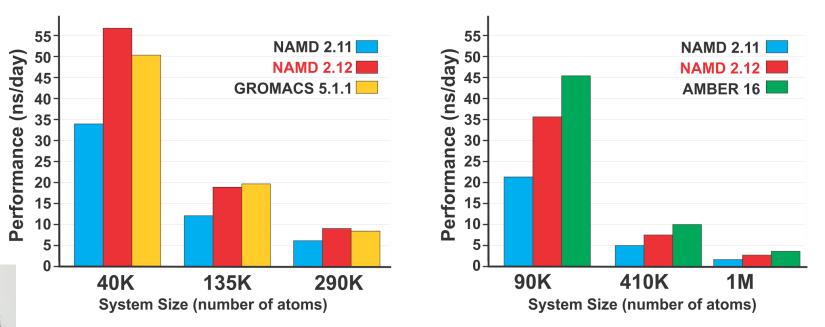
Streaming Kernel



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



Single-Node GPU Performance Optimization

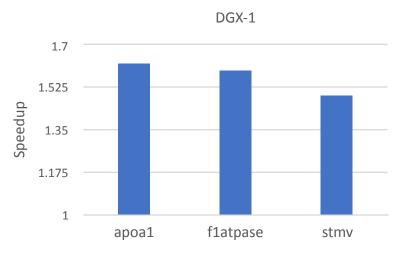


New kernels by **Antti-Pekka Hynninen**, formerly Oak Ridge, **NVIDIA**. Stone, Hynninen, et al., *International Workshop on OpenPOWER for HPC (IWOPH'16)*, 2016.

Described at GTC 2016 S6623 - Advances in NAMD GPU Performance

Coming in NAMD 2.13: Bonded force offloading

- GPU offloading for bonds, angles, dihedrals, impropers, exclusions, and crossterms
- Computation in single precision
- Forces are accumulated in 24.40 fixed point
- Virials are accumulated in 34.30 fixed point
- Code path exists for double precision
 accumulation on Pascal and newer GPUs
- Reduces CPU workload and hence improves performance on GPU-heavy systems



New kernels by Antti-Pekka Hynninen, NVIDIA.

Summit will replace Titan as the OLCF's leadership supercomputer



- Many fewer nodes
- Much more powerful nodes
- Much more memory per node and total system memory
- Faster interconnect
- Much higher bandwidth between CPUs and GPUs
- Much larger and faster file system

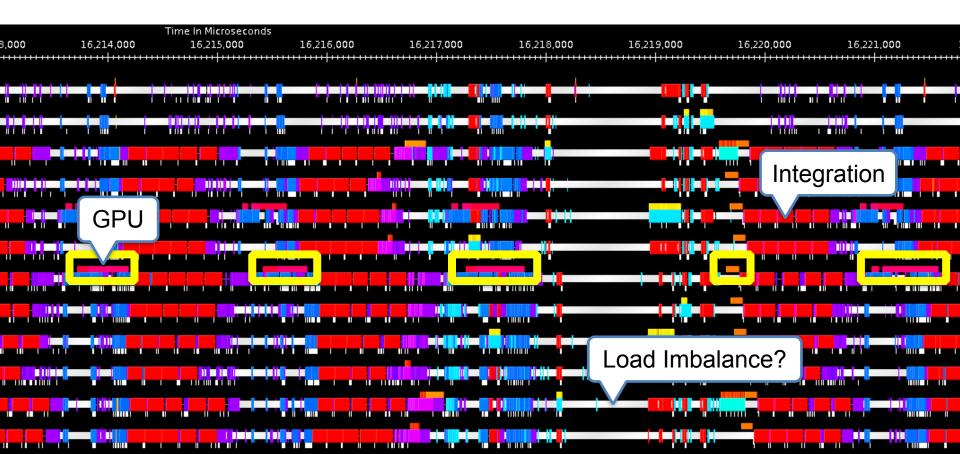
Titan	Summit
Baseline	5-10x Titan
18,688	~4,600
1.4 TF	> 40 TF
32 GB DDR3 + 6 GB GDDR5	512 GB DDR4 + HBM
0	1600 GB
710 TB	>10 PB DDR4 + HBM + Non-volatile
Gemini (6.4 GB/s)	Dual Rail EDR-IB (23 GB/s)
3D Torus	Non-blocking Fat Tree
1 AMD Opteron™ 1 NVIDIA Kepler™	2 IBM POWER9™ 6 NVIDIA Volta™
32 PB, 1 TB/s, Lustre [®]	250 PB, 2.5 TB/s, GPFS™
9 MW	15 MW
	Baseline 18,688 1.4 TF 32 GB DDR3 + 6 GB GDDR5 0 10 10 10 10 10 10 10 10 10 10 10 10 1



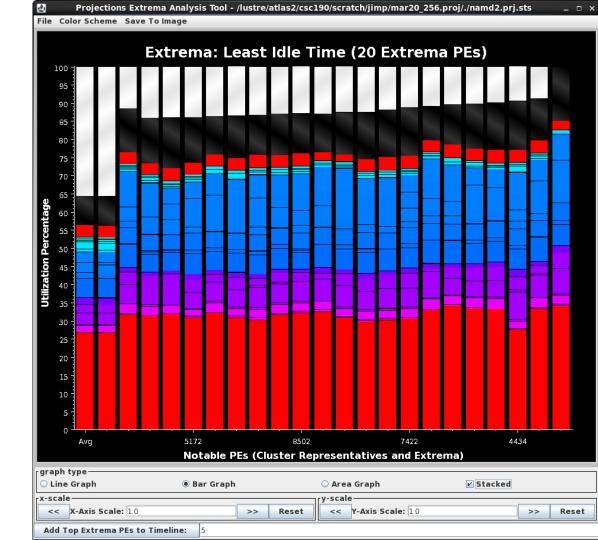
Charm++/NAMD configuration

- IBM PAMI SMP machine layer
 - Initially developed for Blue Gene series
 - No dedicated communication thread
- Single GPU per process (6 processes per node, 6 threads per process)
 - Leaving one core free per resource set seems to reduce noise
 - One core per socket is reserved by jsrun, so 8 unused cores per node
- With thread to core affinity:
 - jsrun -r6 -g1 -c7 namd2 +ignoresharing +ppn 6 +pemap 4-27:4,32-55:4,60-83:4,92-115:4,120-143:4,148-171:4
- Or without (expected to run slower, but sometimes faster):
 - jsrun --bind rs -r6 -g1 -c7 namd2 +ignoresharing +ppn 6

Charm++ Projections tool shows bottleneck



Charm++ *Projections* Extrema Tool Finds Problem PEs



One PE has no idle time! Also, overloaded PEs are all GPU hosts

	16,209,000	16,210,000	16,211,000	16,212,000	16,213,000	16,214,000	Time In Microseconds 16,215,000	16,216,000	16,217,000	16,218,000	16,219,000	16,220,0
	 	•••••	••••••	••••••	••••••		••••••	•••••	*****	•••••	••••••	 .
PE 0 (21, 17)	, <u>, , , , , , , , , , , , , , , , , , </u>	· <mark>-</mark>		── ! ──!			,			* <mark>```</mark> =1•¥ ··· • !!		
PE 1 (21, 17)				÷ II - II - II -				<mark>,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,</mark>		-)-)-(-(
PE 2 (71,60)	•• ••• •••••••••••••••••••••••••••••••						, , , , , , , , , , , , , , , , , , , 			· ···		
PE 3 (57, 48)		,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,										
PE 4 (70, 60)												
PE5 (60, 53)							et e					
PE 6 (77, 65)												
PE 7 (65, 54)	())= () = 		(
(65, 54) PE 8 (60, 53)			•									
PE 9			1									
(59,51) PE10												
(69,59) PE11												
(65, 56) PE 1014			· ·									
(100, 84) PE 4074												
(88, 76) PE 4434												
(88, 75) PE 5046												
(90, 77) PE 5154												
(89, 76)												

Same issue on 512 nodes Now showing all PEs on process

	16,230,000	16,230,400		16,231,200	16,231,600	16,232,000	16,232,400		16,233,200	16,233,600	16,234,000		16,234,800	16,235,200
PE 0			·····									·····		
(22, 17)		╬╬╾┵╬╬╌┶┷╻┿╛	n na star											
PE 1 (21, 16)		••••••••••••••••••••••••••••••••••••••		0.000	<u>0-0-0-0-0</u>) -1				•		• <mark>₽-₩-₩</mark>
PE 2 (62, 51)			در در <mark>از در ا</mark> ست ا		<mark>, </mark>	excer m oke <mark>t</mark> en		·····			- <mark>11 - 12 - 14 - 14 - 14 - 14 - 14 - 14 - </mark>			- 10 <mark>10-000</mark>
PE 3 (69, 58)														
PE 4 (68, 57)									₽₩					, , , , , , , , , , , , , , , , , , ,
PE 5 (62, 54)														
PE 6 (67, 58)													, i	
PE 7 (58, 49)			- 					━━┾═━╤━╡╢						
PE 8 (69, 60)								((); ()=(i)=(i)=(i)=(i)=		
PE 9 (59, 50)									+ (())) +					() - (.
PE 10 (80, 67)					<mark>,,,,,</mark> ,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,			+)) <mark></mark> 0						
PE 11 (59, 50)								── ────!── ₩						
PE 8232 (95, 80)														
PE 8233 (56, 48)							┿╾╾╾┿ <u>╎╏<u>╢╢</u>┥╺┝╹</u>					-))- <mark>-</mark> -))-)-(
PE 8234 (54, 47)				,, .,,								-)() <mark>0</mark>) <u>0</u> -(+-(
PE 8235 (63, 52)														
PE 8236 (56, 48)												••••••••••••••••••••••••••••••••••••••		
PE 8237 (59, 50)														
PE 8238 (65, 55)											aiia <mark>i ii</mark> hijataaliija <mark>i</mark>			
PE 8239 (70-60)		-												

Try removing patches from GPU host PEs Overloaded PEs (256 nodes) are no longer GPU hosts

	Time In Microseconds
	9,004,000 9,005,000 9,012,000 9,012,000 9,012,000 9,014,000 9,016,000 9,016,000 9,022,000 9,024,000 9,026,000 9,028,000 9,030,000 9,032,000 9,034,000 9,036,000 9,040,000 9,042,000 9,044,000 9,044,000 9,046,000 9,000 9,000 9,000 9,000 9,000 9,000 9,000 9,000 9,000 9,000 9,000 9,000 9,000 9,000 9,
	\sim
PE 0 (18, 15)	
(18, 15) PE 1	
(18, 15)	
PE 2	
(80,68)	
PE 3 (65, 55)	
PE 4	<u>⋒⋒⋒⋒⋒⋒⋒⋳⋳⋳⋳⋳⋳⋳⋳⋳⋳⋳⋳⋳⋳⋳⋳⋳⋳⋳⋳⋳⋳⋳⋳⋳⋳⋳⋳⋳⋳</u>
(24, 20)	
PE 5 (68, 59)	
PE 6	
(36, 30)	
PE 7 (76, 66)	
PE 8	
(76, 65)	
PE 9 (71, 63)	
PE 10	
(69, 61)	
PE 11 (75, 65)	
PE 1015	
(92, 79)	
PE 7039	
(90, 79) PE 811	
(89, 77)	
PE 8032	
(89, 77)	
PE 8613 (89, 77)	

•				►
Display Pack Times	🗾 Display Message Sends	🗹 Display Idle Time	🔽 Display User Events	View 793 User Events
e,	Zoom Ratio:	1.0		Reset Zoom

Overloaded PEs still have idle time Now showing all PEs on process

			Time In Microseconds		.000 9.042.000 9.044.000 9.045.000 9.048.000 9.050.000
PE 2 (80, 68)			······································		,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,
PE 3 (65, 55)		,			
PE 4 (24, 20)		<mark>en anaien dau e 2 - a anten en a</mark>			
PE 5 (68, 59)		······································			
PE 6 (36, 30)		, <mark>10-2020, 2020, 102, 102, 102, 102, 102, 102</mark>	()-(<mark>))) ()</mark> ()-())())-()()()-()()()()-())()()()(
PE 7 (76, 66)		P02			
PE 8 (76, 65)		waa			
PE 9 (71, 63)		****			
PE 10 (69, 61)		···			
PE 11 (75, 65)		72		,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	
PE 1014 (42, 35)		. 			
PE 1015 (92, 79) PE 1016					
(80, 71) PE 1017					
(70, 62) PE 1018					
(84, 74) PE 1019		1991 - Hold - Calendary - Hold			
(79, 68) PE 7038					
(37, 31) PE 7039				r ann an 1975 ann an 1977 ann an 1977 ann ann a' 1977 ann ann an 1977 ann	
(90, 79) PE 7040					-
(75, 66) PE 7041					
(72, 63)					
	Display Pack Times	🗹 Display Message Sends	🗹 Display Idle Time	🗹 Display User Events	View 1711 User Events
	ę	Zoom Ratio:	1.0	e,	Reset Zoom

Load imbalance delayed until PME steps 256 nodes zoomed in

	Time In Micro																
	9,025,800	9,027,200	9,027,600	9,028,000	9,028,400	9,028,800	9,029,200	9,029,500	9,030,000	9,030,400	9,030,800	9,031,200	9,031,600	9,032,000	9,032,400	9,032,800	9,03
PE 2								•••••••									<u> </u>
(80, 68) PE 3							111						1 1 ¹				
(65, 55)			أستاس														den -
PE 4 (24, 20)	<u> </u>												~~~~~				
PE 5 (68, 59)									━╹━╹	.				.			
PE 6 (36, 30)							<mark>╔╔┇</mark> ╼╧ <mark>╸</mark> ╼╧╸╬╫		━━₩,₽,₩━							• • • • •	
PE 7 (76, 66)										•							
PE 8 (76, 65)																	
PE 9					•) 												
(71,63) PE10																	
(69,61) PE11																	
(75, 65)																	
PE 1014 (42, 35)	m			— <u>()-()</u> ,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,									<u> -)- - - </u>				
PE 1015 (92, 79)																	
PE 1016 (80, 71)		i an															
PE 1017 (70, 62)	<u> </u>																
PE 1018 (84, 74)		iiiii a a iii aa a iiii															
PE 1019		ر ر در این ا رز ارز ا															
(79, 68) PE 7038																	· · · · ·
(37, 31) PE 7039															i	nn <mark>i i i i</mark>	1 111
(90, 79)				i de la com													
PE 7040 (75, 66)																	
PE 7041 (72, 63)																	╴╻╴
	Display Pack	Times		🗹 Displ	lay Message Sen	ıds		🗹 Display Id	lle Time		✓ Displa	ay User Events			🗌 View 1711 Use	er Events	
	Q			Zo	om Ratio:		7.0					R			Reset Zoon	n	

512 nodes similarly improved Showing overloaded PEs

Time In Microseconds
14.808.000 14.810.000 14.811.000 14.812.000 14.813.000 14.814.000 14.815.000 14.815.000 14.816.000 14.819.000 14.819.000 14.822.000 14.823.000 14.824.000 14.825.000 1

PE 0 (19,16)	<u>, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997</u>	<u> </u>	<u>,</u>
PE1 (19,15)	~}~;~;;;;?????????????		`````````````````````````````````````
PE 2 (69, 57)	*************************************		de la superior de la superior ente de la superiorente de la s
PE 3 (63, 52)			
PE 4 (29, 24)	₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩	יד ארביינוער איזיידער ביאידער ארביי איזי ארבייער איז ארביין ארביין ארביין ארביין ארביין ארביין ארביין ארביין אר	<u>جا کرت کا این مراجع میں اور میں میں اور اور میں اور میں اور میں میں میں میں میں مواجع کر اور میں مراجع میں م</u> رک
PE 5 (63, 53)			
PE 8 (72, 59)			
PE 11 (71,59)			······································
PE 12004 (84, 71)	moli relation de la section de la section de la section		
PE 559 (84, 70)			
PE 6115 (83, 69)			
PE 7396 (83, 70)			
PE 2074 (82, 69)			

•				
Display Pack Times	🖌 Display Message Sends	🗹 Display Idle Time	🗾 Display User Events	View 572 User Events
e,	Zoom Ratio:	1.0		Reset Zoom

512 nodes similarly improved Now showing all PEs on process

Time In Microseconds
14.809.000 14.810.000 14.812.000 14.812.000 14.813.000 14.814.000 14.815.000 14.815.000 14.816.000 14.816.000 14.816.000 14.816.000 14.820.000 14.820.000 14.822.000 14.823.000 14.825.000 1

PE 2					
(69, 57) PE 3					
(63, 52) PE 4					
(29, 24)			,	₩ <u></u>	
PE 5 (63, 53)					<u>a in 1966 in 1966 and an 1966 a</u>
PE 6 (40, 33)		and and a second state of the s	میں براہ است ارباد و برای اور کر در در در سرکر کا در ارباد کر اور		an a
PE 7 (58, 49)		· · · · · · · · · · · · · · · · · · ·			
PE 8					
(72, 59) PE 9					
(67, 56) PE 10					
(73, 63)					
PE 11 (71,59)			-0.4	in a state of the second s	
PE 558 (37, 31)	~~~~ ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	and the second secon			
PE 559 (84, 70)					
PE 560					
(72,60) PE561					
(67, 57) PE 562					
(58, 50)			,		
PE 563 (61, 52)		<u>رداری از میشنان فروندی وروند از موارد ا</u>		-04	
PE 12000 (44, 38)			ار دین پر بسیان اور <mark>اور اور اور اور</mark> اور		
PE 12001 (79, 67)			, a particular de la factoria de la factoria de la competencia de la competencia de la competencia de la compe		
PE 12002			واخدا الأدر والمتحد المتحد والألب المتكر والمتحد والمتحد والمتحد والمتحد والمتحد والمتحد		
(59, 51) PE 12003					
(69, 59)					
	Display Pack Times	🗹 Display Message Sends	🗹 Display Idle Time	🗹 Display User Events	View 1251 User Events
	e,	Zoom Ratio:	1.0	e,	Reset Zoom

GPU overlaps with communication only 512 nodes zoomed in

Time In Microseconds 14,817,200 14,817,400 14,817,600 14,817,800 14,818,000 14,818,200 14,818,400 14,818,600 14,818,800 14,819,000 14,819,200 14,819,400 14816400 14816600 14816800 14817000 14 819 800 14 820 600 14 820 800 14 821 000 14 821 2 (69, 57) (63.52) (29, 24) (63, 53) (40, 33) a a sha a (58, 49) (72, 59) (67.56) (73, 63) (71, 59) PE 558 (37, 31) PE 559 (84, 70) PE 560 (72, 60) PE 561 (67, 57) PE 562 (58, 50) PE 563 (61, 52) PE 12000 (44, 38) a community PE 12001 (79, 67) PE 12002 (59.51) PE 12003 (69, 59) Display Pack Times Display Message Sends ✓ Display Idle Time Display User Events View 1251 User Events Q €. Zoom Ratio: 4.5 Reset Zoom

Notes on Benchmarks

- All results are early and preliminary.
- We've had access for less than two months.
- Acceptance is not until this summer.
- Only 1/4 of the nodes are available.
- Installation and software testing continue.
- New platforms always have issues.
- The Volta GPUs are really fast!



We're ignoring the noise problem for now

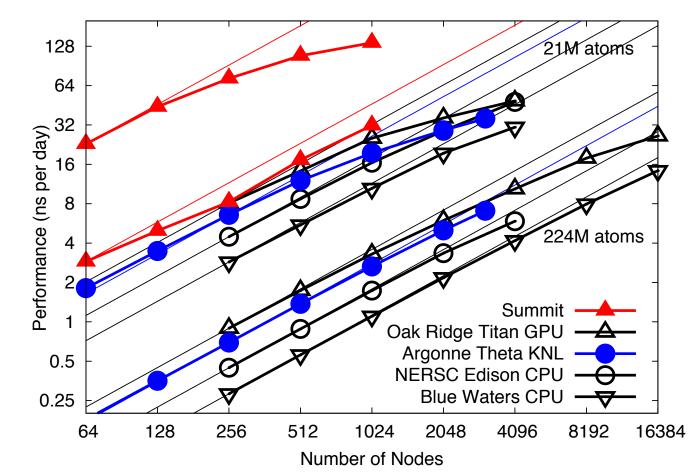
		14.200.000	14.240.000	14.280.000	14.320.000	14.360.000	14.400.000	Time In Microseco 14.440.000	14,480,000	14.520.000	14.560.000	14.600.000	14.640.000	14.680.000
		H			••••••	••••••	·····	•••••		•••••		••••••		
r=121 (41.38) distribution of the bit of t		nisin King b		in the initial		1 1		- minutioni						
(a) 100 100 100 100 100 100 100 100 100 10														
PE 107 PE														
PE1014 Image: State of the state of t						it.								
195.32/2 195.32/2 <td< td=""><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td>1011101010</td><td></td><td></td><td></td><td></td><td></td><td></td></td<>								1011101010						
PE 2993 abapted bill bill bill bill bill bill bill bil						1								
P2 2003 ab babada dar kontra - abada dar kontre - abada dar kontre - abada dar kontra - abada dar ko						<u>i</u>								
PE3386 Image: Person of the second of th	PE 2889			1 111 101 111		h		والمتحاطين والمتحاطي				111		taiten an de
PE 3070 PE 4074 PE 4074 PE 4074 PE 5072 PE 5074 PE 5074 <td></td> <td></td> <td></td> <td></td> <td></td> <td>÷.</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td>						÷.								
PE 4074 PE 4382 PE 4434 PE 5046 PE 4434 PE 5046)I								
PE4362 (33, 28) Image: Constraint of the second of the s						<u>11</u>								
(33,28) (34,434) (34,29)								10.11010.000				1011		
(3, 29) (3, 29)	(33, 28)													
PE5046 (34, 29) Image: Image						î								
PE5082 (34,29) Image: Constraint of the state of the	PE 5046											1.8		
PE5154 (34,29) Image: Comparison of the state of the														
PESIDA (33,29) Image:						1								
PE5172 (33,28) PE5592 (43,39) PE6918 (33,28)						ii.								
(33,28) PE5592 (43,39) PE6518 (33,28) (33,28)								TELEVISION OF A DESCRIPTION OF A DESCRIP				1111		
	(33, 28)													
		di-110,010				1		U(10100)				00		
	PE 6918													
		o cantana				1								

80 ms

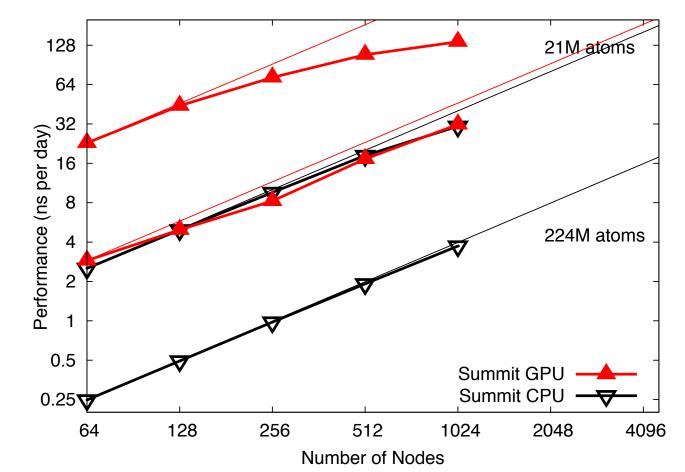


Biomedical Technology Research Center for Macromolecular Modeling and Bioinformatics Beckman Institute, University of Illinois at Urbana-Champaign - www.ks.uiuc.edu

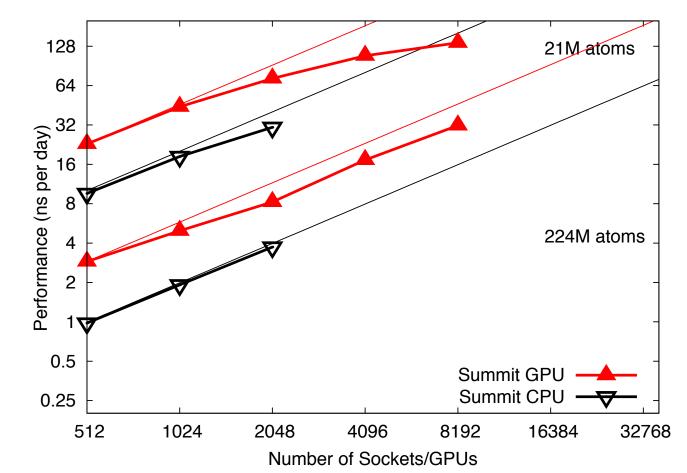
Comparison for large benchmarks



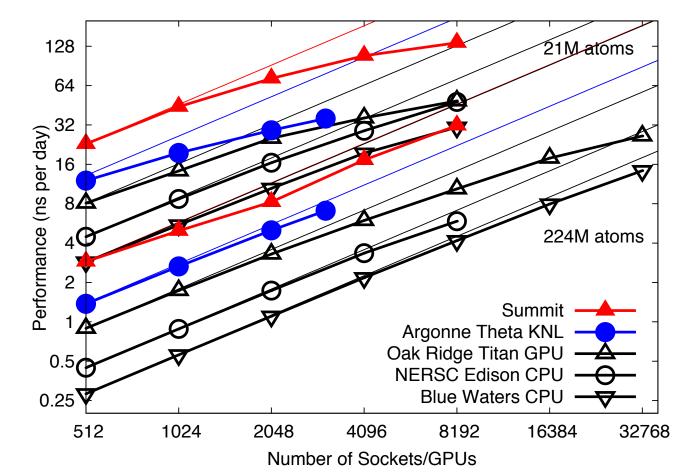
Comparison for large benchmarks



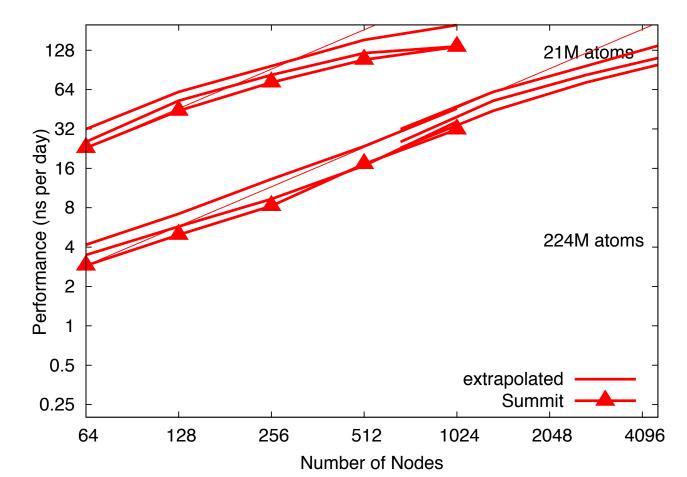
"Fair" comparison for large benchmarks



"Fair" comparison for large benchmarks



"Fix" problems, extrapolate from 21M atom results



Conclusions and Future Work

- Summit represents a new era in GPU acceleration
 - The CPU will be the bottleneck for many codes
 - Optimizing/vectorizing/parallelizing on the CPU not enough
 - Offload everything practical to the GPUs
- Worry about optimizing the CUDA code last
 - Stage/stream data to reduce CPU/network bottlenecks
- A supercomputer is not just a large cluster
 - IBM knows this (Blue Gene series), Summit should scale well

Acknowledgments

Antti-Pekka Hynninen & Ke Li, NVIDIA Sameer Kumar & Bilge Acun, IBM Tjerk Straatsma, OLCF William Kramer, NCSA Alexander Bobyr & Michael Brown, Intel Abhi Singharoy, ASU



NIH Center for Macromolecular Modeling and Bioinformatics University of Illinois at Urbana-Champaign









