D EXAMPLE SCIENCE AREA DRIVERS AND REQUIREMENTS OVER TIME

2014 Best of Breed based on Blue Waters Science and Engineering teams.

Science Categories	Goal	Fidelity	Simulated Time Duration	Number of Simulation Problems	Data Integration and sharing	Algorithm and Workflow re-engineering	Long Range
Astronomy, Astrophysics and Space Science **Best of Breed**:	First 3D multi-physics general-relativistic magnetohydrodynamics simulations of core-collapse supernovae. Detailed microphysics: nuclear equation of state, neutrino heating/cooling.	~400 3D grid functions on 1 billion mesh zones, finely resolving magnetorotationally-driven outflows out to 2000 km. Simulation volume: 10^4 km	0.5 seconds		Yes	In progress for GPU-CPU hybrid operation.	Yes
	Simulating a full star (radius ~10^3 km); initial ignition point (r ~ 2 km); effective resolution $36{,}864^3$ (135 m/zone)[m]; Full 3D general relativity; GR magneto-	$1{,}536^3$ cells for 5.7 million time steps[n]Simulating a full star (radius ~10^3 km); initial ignition point (r ~ 2 km); effective resolution $36{,}864^3$ (135 m/zone)[o]; Full 3D general	1 day	MHD box 5200×2880×2800 astronomical units (AU) cubed;		Adaptive Mesh Refinement 5- 10 levels of AMR –	

[m] Chris Malone's *Blue Waters 2014 Symposium* Presentation (http://bluewaters.ncsa.illinois.edu/symposium-2014-schedule-descriptions)

[n] Paul Woodward's *Blue Waters 2014 Symposium* Presentation (http://bluewaters.ncsa.illinois.edu/symposium-2014-schedule-descriptions)

[o] Chris Malone's *Blue Waters 2014 Symposium* Presentation (http://bluewaters.ncsa.illinois.edu/symposium-2014-schedule-descriptions)

Science Categories	Goal	Fidelity	Simulated Time Duration	Number of Simulation Problems	Data Integration and sharing	Algorithm and Workflow re-engineering	Long Range
	hydrodynamics with detailed nuclear equation of state and neutrino heating/cooling	relativity; GR magneto-hydrodynamics with detailed nuclear equation of state and neutrino heating/cooling					
	Several Hour of simulated Heliosphere time with full kinetics	Helio Science – fully kinetics – 10^{10} cells, $4x10^{12}$ particles; Large Scale Hybrid Kinetic – $1.7x10^{10}$ cells, $2x10^{12}$ particles.[p] Coupled continuum and kinetic equations – 10^{10} particles; local effective resolution on the MHD grid $5{,}000^3$ for a very large computational box.	Several Hours			Hybrid parallelization, optimization of the source term calculations in Monte Carlo simulations, improved load balancing for Monte Carlo algorithms	
Biophysics and Biology **Best of Breed**:	All Atom Molecular Dynamics – e.g Virus Interaction, Cell Organelles (Ribsomes, Chromatiphore)	Best of Breed – 100M Atoms[r]	Microseconds	2 to 4	Yes	Higher Order PME interpolation	Yes

[p] Vadim Roytershteyn's *Blue Waters 2014 Symposium* Presentation (http://bluewaters.ncsa.illinois.edu/symposium-2014-schedule-descriptions)

[r] Klaus Schulten's *Blue Waters 2014 Symposium* Presentation (http://bluewaters.ncsa.illinois.edu/symposium-2014-schedule-descriptions)

Science Categories	Goal	Fidelity	Simulated Time Duration	Number of Simulation Problems	Data Integration and sharing	Algorithm and Workflow re-engineering	Long Range
	Ultra Coarse Grain Simulations – e.g. Virus interactions Multiple Copy Simulations[q]	1M Atoms		1		Ultra Coarse Grain	
Chemistry **Best of Breed**:		Systematic, predictive, but extremely expensive ("ab initio") methods developed over the	Interested in static properties of materials			The computational cost of density functional theory (DFT) is routine for many types of solids and is so enormous it is necessary to re-engineer the algorithms.	
Climate and Atmospheric Sciences **Best of Breed**:		1/4° Atm, 1° ocean[s] Radiative heating rates for real	~250 years	1-4 for climate, 10 for others	Yes		

[q] The reference to multiple copy simulations must be Phillips, Roux and Schulten

[s] Don Weubbles' *Blue Waters 2014 Symposium* Presentation (http://bluewaters.ncsa.illinois.edu/symposium-2014-schedule-descriptions)

Science Categories	Goal	Fidelity	Simulated Time Duration	Number of Simulation Problems	Data Integration and sharing	Algorithm and Workflow re-engineering	Long Range
		(i.e., 3D) clouds and their impact on dynamics.[t]					
Geophysics/Seismic **Best of Breed**:	Dynamic Rupture Simulation to understand physics of Earthquake Rupture Process	5-Hz simulation in small, fault-oriented, volume. 150kmx100kmx40km 82K time steps	100 sec	5-10 runs 37-billion mesh elements	Yes	Use Dynamic rupture simulation to create inputs for high frequency wave propagations	Yes
	High-frequency Scenario Earthquake Simulation	Wave propagation run: 2-Hz, 810x405x85km, 40m spacing, Min. S-wave Velocity 400m/s, 160K time steps	250 sec	2-4 runs 435 billion mesh points		Heterogeneous computing on Blue Waters	
	Regional Probabilistic Seismic Hazard Analysis	1144 seismic hazard sites at 0.5 Hz UCERF 2 700,000 rupture variations used to simulate 500 million 2-component seismograms[u]		1.2 billion points/mesh per wave propagation simulation		Empirical Ground Motion Probability Models replaced with physics-based ground motion models;	

[t] Larry Di Girolamo *Blue Waters 2014 Symposium* Presentation (http://bluewaters.ncsa.illinois.edu/symposium-2014-schedule-descriptions)

[u] Tom Jordan's *Blue Waters 2014 Symposium* Presentation (http://bluewaters.ncsa.illinois.edu/symposium-2014-schedule-descriptions)

Science Categories	Goal	Fidelity	Simulated Time Duration	Number of Simulation Problems	Data Integration and sharing	Algorithm and Workflow re-engineering	Long Range
						Seismic Reciprocity. 100 million jobs 12 TB data products	
Fluids and Turbulence **Best of Breed**:	Turbulent Flows Ice nucleation in turbulent domains	$8{,}192^3$ cells[v] (over 0.5 trillion grid points) 1M Atoms droplet	5-7 large-eddy turnover times	1 (over 3 years)	Yes - Sharing reference simulation data sets	Reducing Communication costs and implementing many-core methods	
Materials Science **Best of Breed**:	Transistor Materials Condensed Matter/ Large Scale Quantum Simulations[w]	100 Million atoms; 1, 2, 3D; any crystal; Spin and Classical Multi-Physics; Transport; Strain MVFF[y], problem specific assumptions 10,000 atoms				DFT Monte Carlo	Yes

[v] P.K. Yeung’s *Blue Waters 2014 Symposium* Presentation (http://bluewaters.ncsa.illinois.edu/symposium-2014-schedule-descriptions)

Science Categories	Goal	Fidelity	Simulated Time Duration	Number of Simulation Problems	Data Integration and sharing	Algorithm and Workflow re-engineering	Long Range
	Quantum Monte Carlo[x]	1,000 atoms					
Particle Physics **Best of Breed**:		HISQ gauge –$96^3 \times 256$ Clover quark propagators - $32^3 \times 256$[z]		600 Configurations 485 configurations	Yes		
Social Science, GIS, Economics **Best of Breed**:		260,000 agents – US Scale[aa]	1 to 2 Months	1 scenario	Yes	Agent	Yes

[w] Jerry Bernholc

[y] Gerhard Klimeck's *Blue Waters 2014 Symposium* Presentation (http://bluewaters.ncsa.illinois.edu/symposium-2014-schedule-descriptions)

[x] e.g David Ceperly's work

[z] Steve Gottleib's *Blue Waters 2014 Symposium* Presentation (http://bluewaters.ncsa.illinois.edu/symposium-2014-schedule-descriptions)

[aa] Keith Bisset's *Blue Waters 2014 Symposium* Presentation (http://bluewaters.ncsa.illinois.edu/symposium-2014-schedule-descriptions)

2017-2018 Best of Breed and Community Normal Science Drivers

Science Categories	Goal	Increased Fidelity	Increased Simulated Time	Number of Simulation Problems	Data Integration and sharing	Algorithm and Workflow re-engineering
Astronomy, Astrophysics and Space Science **Best of Breed**:	Star Ignition	Simulations of local small-scale processes (magnetic reconnection, plasma instabilities, turbulence, etc.) in the framework of large-scale global MHD simulations.	10x increase – 1 day of simulated time	10x simulations	Yes	Optimized discrete event simulations; code-code coupling; semi-implicit 3D PIC with high order particles; development of accurate subgrid models
	Helio Science – Turbulence and global modeling	Currently, small- and micro-scale phenomena are addressed only locally, which makes difficult to predict their effect on the global solution.	Requires 10^4 more computing time than in the best contemporaneous simulations.	Total volume is not increasing, but the grid is 10 times finer in all directions locally.		AMR and load balancing improvements; I/O issues for very large data files (40-50 TB).
Community Std:						
Observational Support:		Voyager and Interplanetary Boundary Explorer missions.				
Biophysics and						

Science Categories	Goal	Increased Fidelity	Increased Simulated Time	Number of Simulation Problems	Data Integration and sharing	Algorithm and Workflow re-engineering
Biology **Best of Breed**:	All Atom Molecular Dynamics Ultra Coarse Grain Simulations – e.g. Virus interactions	Multiple Cellular Structures- 200-300M all atom Polarizable force fields and O(N) long range dispersion forces	10 microseconds			O(N) electrostatis
Community Std:	Multiple Copy Simulations[bb]	10M Atoms	Microseconds	10s of problems		Streaming Data Flow Analysis
Observational Support:		Cellular Structures100M all atom Large all to all comparisons				
Chemistry **Best of Breed**: **Community Std:**						
Climate and Atmospheric Sciences ***Best of Breed***:		Climate 1/8° (12 km)	250-500 years	3 to 4	Yes	Atmospheric chemical tracers,

[bb] see Roux and Tajkorshid

Science Categories	Goal	Increased Fidelity	Increased Simulated Time	Number of Simulation Problems	Data Integration and sharing	Algorithm and Workflow re-engineering
		Weather .1 km IPCC - Climate 1/8° (12 km) Climate 1/4° (25 km	 500 years 1,000 years	 3-4 20-30		indirect radiative balance, nested grids for regional impacts
Geophysics **Best of Breed**:	Physics of Earthquake Rupture Process High Frequency Scenario Earthquake Simulation	Dynamic rupture simulations: 100x50x50km, 2.5m spacing, 16 trillion mesh points, 200K time steps Wave propagation simulations: 5-Hz, 810x405x85km, 10m spacing, min. S-wave Velocity 250m/s, 640K time steps	150 secs 250 secs	2-4 32 trillion points/mesh 80 billion points/mesh per wave propagation simulation	Yes	Scalability and load rebalancing of discontinuous mesh algorithm. Wave propagation simulation is run longer than dynamic rupture simulation to allow waves to propagate across full simulation volume for wave propagation simulation
Community Std:	Regional Probabilistic Seismic Hazard Analysis	3500 California statewide seismic hazard sites at 2.0 Hz using UCERF3 earthquake rupture forecast An estimated 23 million rupture variations used to simulate 50 billion 3-component seismograms Improved physics (frequency-dependent attenuation, fault roughness, near-fault plasticity, soil non-linearities, near-surface	250 secs	3,500 hazard curve calculations		9.5 billion jobs 7.2 PB data products

Science Categories	Goal	Increased Fidelity	Increased Simulated Time	Number of Simulation Problems	Data Integration and sharing	Algorithm and Workflow re-engineering
		heterogeneities)				
Fluids and Turbulence **Best of Breed**: **Community Std:**		$8{,}192^3$ simulations cells $4{,}096^3$ simulations cells	Increased	4-5 10-20	Yes	Increased complexity, multiple parameters
Materials Science **Best of Breed**: **Community Std:**	Systematic, predictive ab-initio methods					Methods for stochastic formulation and algorithms Fragment methods
Particle Physics **Best of Breed**: **Community Std:**	Fully dynamical electromagnetic effects Calculation of physical quantities of above configurations	HISQ gauge $-128^3\times256$ Similar increase for Clover quark propagators and DFW actions (14*3) x increase in computational requirements	10x to match precision of latest experimental devices at Fermi Lab	1000 configurations		Inclusion of electromagnetic and isospin breaking effects Improved inversion of sparse matrix

Science Categories	Goal	Increased Fidelity	Increased Simulated Time	Number of Simulation Problems	Data Integration and sharing	Algorithm and Workflow re-engineering
Social Science, GIS, Economics						
Best of Breed:		1,500,000,000 agents + increased agent sophistication – largest country	1 to 2 Months	1 to 2 scenarios	Yes	Dynamic Workload Balancing for scale
Community Std:		260,000,000 agents - US scale	1 to 2 Months	100 scenarios	Yes	

2020-2021 Best of Breed and Community Normal Science Drivers

Science Categories	Goal	Increased Fidelity	Increased Simulated Time	Number of Simulation Problems	Data Integration and sharing	Algorithm and Workflow re-engineering
Astrophysics and Space Science **Best of Breed**: **Community Std:**	Helio Science	Numerical simulation of the turbulent, multi-ion solar wind flow from the solar surface to the Earth's location and its further interaction with the interstellar medium. This will be necessary for the Solar Probe Plus mission to be launched by NASA in 2018. Currently, only some pieces of this approach are implemented, sometimes only locally	The physical times to be covered do not increase, but the grids will be larger, to cover different length scales involved in the simulation.	The physical volume is the same. The problem will be in the division of the computational volume into sub-volumes to make the simulation feasible.	Yes	Using new supercomputing architecture will be essential to implement this challenge because of the additional physics to be added.
Biophysics and Biology **Best of Breed**: **Community Std:**	All Atom MD – e.g virus-cellular interactions	1-10 billion-atom virus-cellular interactions Multiple cellular structures – 200 M Atoms	1-10 Milliseconds Microseconds	2-4 10s to 100s	Yes	Full O(N) quantum mechanics methods for 1M atoms

Science Categories	Goal	Increased Fidelity	Increased Simulated Time	Number of Simulation Problems	Data Integration and sharing	Algorithm and Workflow re-engineering
Chemistry **Best of Breed**: **Community Std:**						
Climate and Atmospheric Sciences **Best of Breed**:	Properly calculate radiative heating rates dynamically within weather and climate models	Climate – 1/16° (6 km) 1/8° (12 km) seasonal prediction of water resources, ultra fine-grain scale to tropical cyclones)	250 years 1,000 years	3-4 20-30	Yes	Nested grids and AMR
Community Std:	Remote sensing problems for retrieving cloud and aerosol microphysical properties in the face of 3D radiative transfer Water Cycles, Ice/Cloud integration/Regional Severe Weather	Weather Prediction	1 day 15 day	Many Regions		

Science Categories	Goal	Increased Fidelity	Increased Simulated Time	Number of Simulation Problems	Data Integration and sharing	Algorithm and Workflow re-engineering
	Prediction for Climate/{Carbon, Methane, Nitrogen} Cycles[cc] Severe Weather – 10m resolution					
Geophysics **Best of Breed**:	Physics of Earthquake Rupture Process	Dynamic rupture simulation: 100x50x50km, 1m spacing, 0.25 quadrillion mesh points, 1 million time steps	250 secs	2-4 runs 1.2 quadrillion points/mesh	Yes	Adopting high precision AMR code for new programming model, load balancing, in-situ analytics, fault tolerance
	High Frequency Scenario Earthquake Simulation	Wave propagation simulation: 10-Hz, 810x405x85km, 3m spacing, min. S-wave Velocity 150m/s, 2.56 million time step	250 secs	2-4 runs 1.2 trillion points/mesh per wave propagation simulation		
Community Std:	Regional Probabilistic Seismic Hazard	20,000 California statewide seismic hazard sites at 5.0 Hz	250 secs	20,000 hazard curve calculations		54 billion jobs 102 PB data products

[cc] http://extremecomputing.labworks.org/climate/reports/ClimateReport_6-24-09.pdf

Science Categories	Goal	Increased Fidelity	Increased Simulated Time	Number of Simulation Problems	Data Integration and sharing	Algorithm and Workflow re-engineering
	Analysis	using UCERF3 earthquake rupture forecast 23 million rupture variations used to simulate 285 billion 3-component seismograms Dynamic rupture models				
Fluid Dynamics **Best of Breed:** **Community Std:**	Ice nucleation in turbulent domains	$16{,}384^3$ simulations cells Realistic droplet – 1 Trillion atoms				Long term strategy to keep multi-Petabyte data collections LES and DNS integration
Materials Science **Best of Breed**:	Multi-dimensions maps for Material Design					
Particle Physics **Best of Breed**:	Increase in 10-100x to keep pace with planned experimental fidelity	128^3*256 HISQ gauge configurations with physical quark masses and lattice spacing		1000 configurations		

Science Categories	Goal	Increased Fidelity	Increased Simulated Time	Number of Simulation Problems	Data Integration and sharing	Algorithm and Workflow re-engineering
Community Std:		Similar increase for Clover quark propagators and DFW actions (16*3)x increase in computational needs				
Social Science, GIS, Economics **Best of Breed**: **Community Std:**		10,000,000,000 agents + increased agent sophistication – world wide 1,000,000,000 agents	1 to 2 Months 1 to 2 Months	1 to 2 scenarios 100 scenarios	Yes	Dynamic Workload Balancing for scale