Early Science at the Argonne Leadership Computing Facility

Paul Messina

Director of Science

Argonne Leadership Computing Facility

and scores of early science participants and ALCF staff

Outline

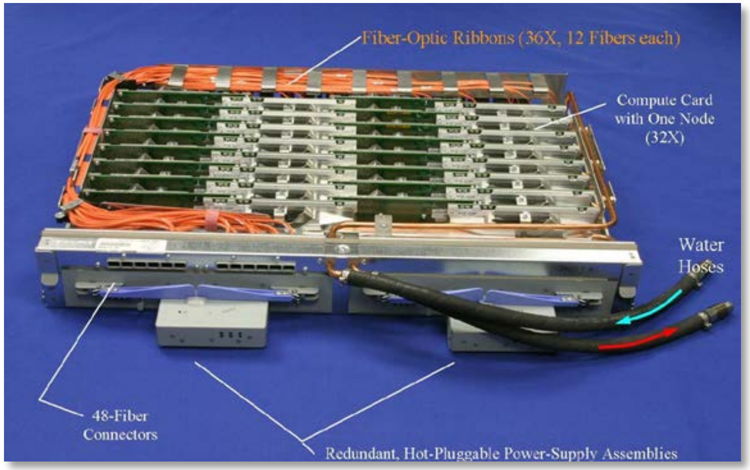
- **Overview of Mira and of BG/Q hardware, programming models**
- **BG/Q Tools and Libraries Project**
- **Status of Mira, T&D system**
- **Mira application benchmarks**
- **Overview of ESP and BG/Q Tools and Libraries Project**
- **Selected applications**
  - Performance, enhancements, insights
- **Summary**

Production ALCF Hardware

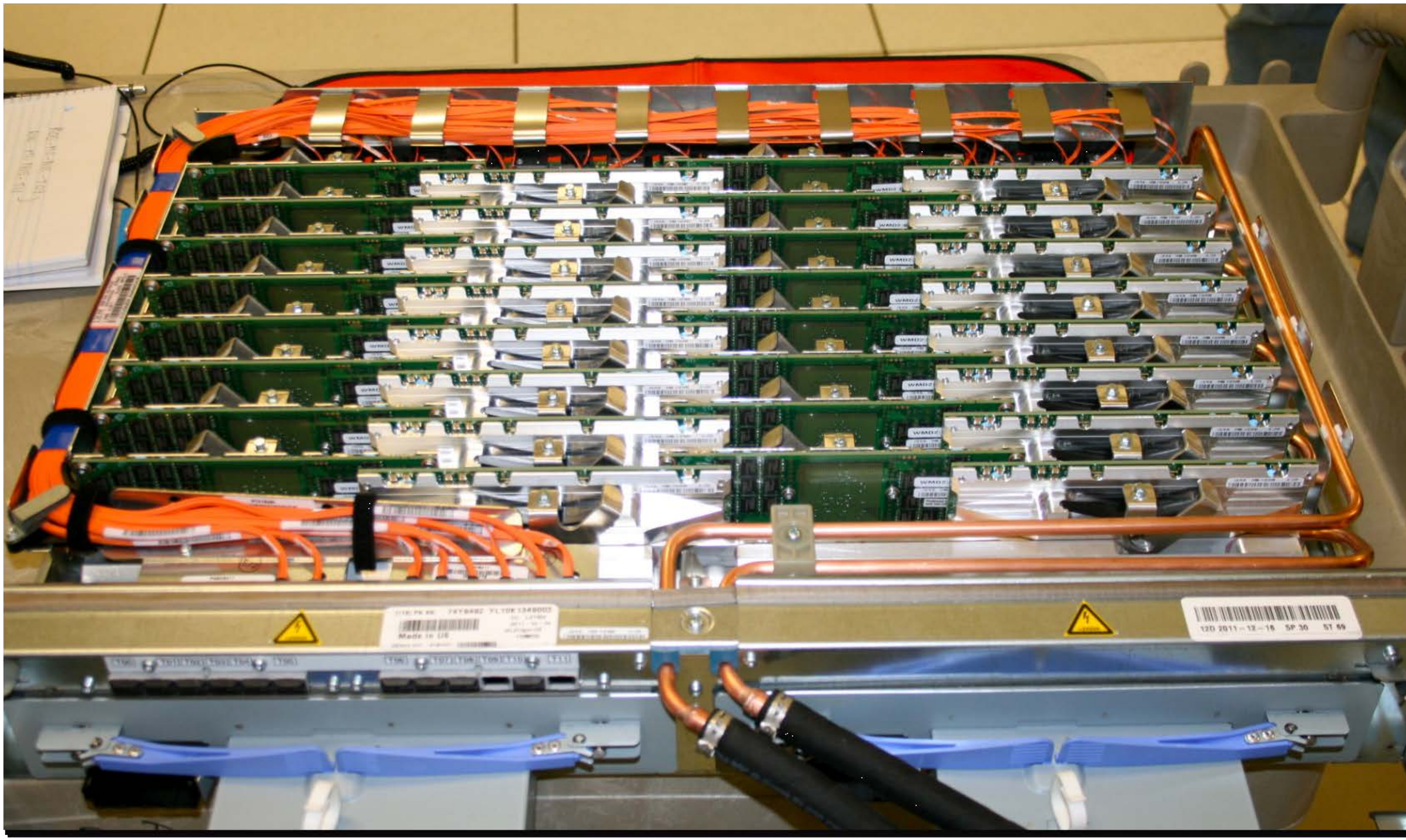
- ***Intrepid* – IBM Blue Gene/P**
  - 40K nodes / 160K PPC cores
    - 40 racks of 1024 nodes
  - 80 terabytes of memory
  - Peak speed: 557 teraFLOPS
  - 6 PB of disk
- **Node**
  - PowerPC 450, 850 Mhz
    - 4 cores
    - 1 HW thread/core
    - 2 GB memory
    - Double-hummer
    - Peak 13.6 GFLOPS
- **Max hardware parallelism:**
  - 163,840

New ALCF Hardware: *Mira*

- ***Mira* – IBM Blue Gene/Q system**
  - 48K nodes, 768K cores
    - 48 racks of 1024 nodes
  - 786 TB of memory
  - 10 petaFLOPS
  - 35 PB of disk
  - Water cooled
- **#3 Top500, #1 Graph 500, #1 Green 500 (tied), all June 2012**
- **Node**
  - PowerPC A2, 1.6 GHz
    - 16 cores
    - 4 HW thread/core
    - 16 GB memory
    - QPX, dual-instruction issue
    - Peak 205 GFLOPS
- **Max hardware parallelism:**
  - 3,145,728

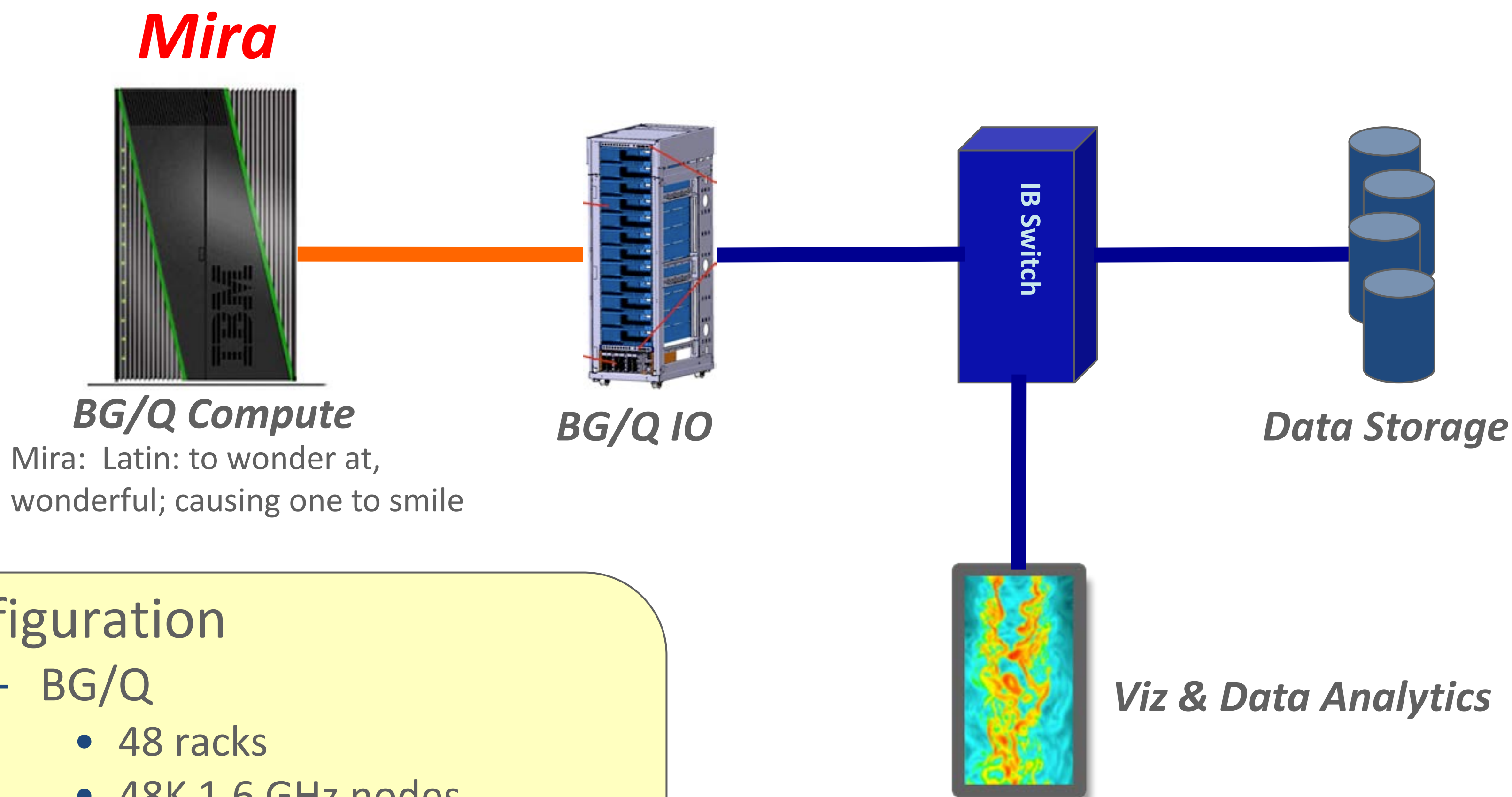



CP
50
45




IBM
MIRA
Blue Gene Q supercomputer

Mira: A 10PF Computational Science Platform



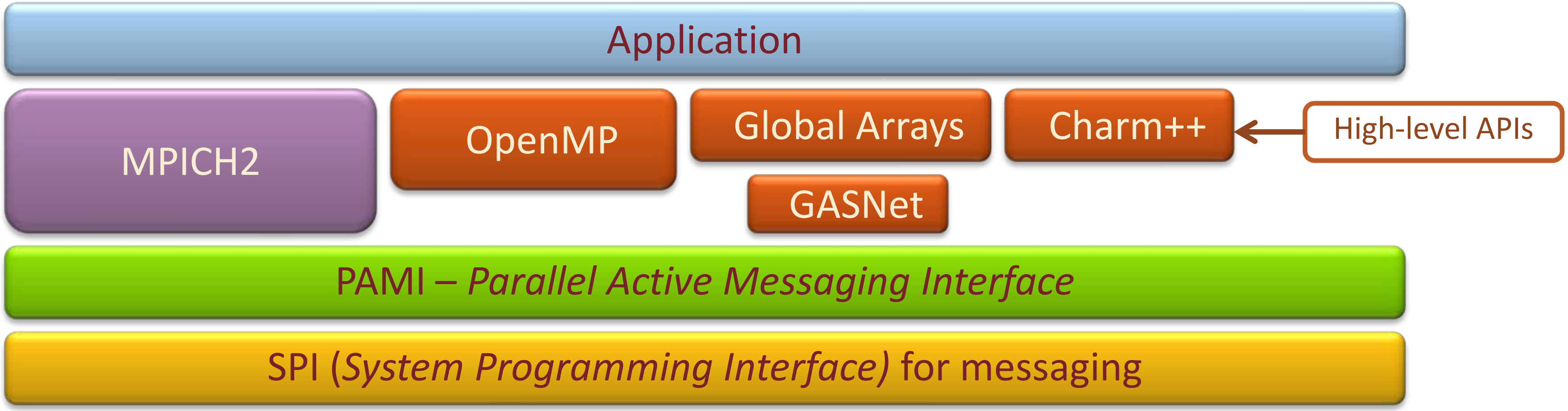
Configuration
- BG/Q
  - 48 racks
  - 48K 1.6 GHz nodes
  - 768K cores & 786TB RAM
  - 384 I/O nodes
- Storage
  - 240 GB/s, 35 PB

Overview of BG/Q vs. BG/P

Design Parameters	BG/P	BG/Q	Improvement
Cores / Node	4	16	4x
Clock Speed (GHz)	0.85	1.6	1.9x
Flop / Clock / Core	4	8	2x
Nodes / Rack	1,024	1,024	--
RAM / core (GB)	0.5	1	2x
Flops / Node (GF)	13.6	204.8	15x
Memory BW/Node (GB/sec)	13.6	42.6	3x
Latency (MPI zero-length, nearest-neighbor node)	2.6 μs	2.2 μs	~15% less
Latency (MPI 64 KB message)	181μs	48 μs	~74% less
Network Interconnect	3D torus	5D torus	Smaller diameter
Concurrency / Rack	4,096	65,536	16x
GFlops/Watt	0.77	2.10	3x

Programming and Running on BG/Q

- **MPI**
- **Threads: OpenMP, PTHREADS**
- **QPX intrinsics: vec_ld, vec_add, vec_madd, ….**
- **Topology interfaces**
  - E.g. MPIX_* functions
- **Run modes: combinations of**
  - MPI ranks/node = {1,2,4,…,64}
  - Threads/node = {1,2,4,…,64}



BG/P applications should run, *unchanged,* on BG/Q — *faster*

Ten Mira Science Application Benchmarks used for acceptance

- Applications cannot be manually tuned; only compiler optimizations are allowed.
- 3 of the applications are threaded – i.e., use both OpenMP and MPI (GTC, GFMC, GFDL).
- The remainder are 100% MPI applications (DNS3D, FLASH, GPAW, LS3DF, MILC, NAMD & NEK 5000).
- For 100% MPI applications, we tested multiple MPI ranks per core (max of 4 ranks per core).
- For MPI + OpenMP applications, we tested 1 MPI rank per core and multiple OpenMP threads per core (max of 4 threads per core)

12

Node/node comparison details

	BG/P	BG/Q	RATIO
Cores / Node	4	16	4x
Clock Speed (GHz)	0.85	1.6	1.9x
Flop / Clock / Core	4	8	2x
Flops/core	3.4 GF	12.8 GF	3.8x
Flops/node	13.6	204.8	15.1x
Nodes / Rack	1,024	1,024	1x
Flops / Rack	13.9 TF	210 TF	15.1x

Mira Science Applications

BG/P version as is on BG/Q

Apps	BQ/P Ratio	Comments
DNS3D	11.8	2048^3 grid, 16K cores, 64 ranks/node
FLASH	5.9 (9.1)	rtflame, 2K cores, 64 ranks/node rtflame, 16K cores, 8 ranks/node, 8 threads/rank, no MPI-IO
GFMC	10.5	c12-test, 2K cores, 8 ranks/node, 8 thrds/rank
GTC	10.8	M0360, 8K cores, 16 ranks/node, 4 thrds/rank
GFDL	11.9	Atm, 2K cores, 16 ranks/node, 4 thrds/rank
MILC	6.1	32^3x64 lattice, 2K cores, 64 ranks/node, no QPX
NEK	7.3	med case, 1K cores, 32 ranks/node, no QPX
NAMD	9.7	ATPase bmk, 2K cores, 16 ranks/node
GPAW	7.6	Au_bulk5x5x5, 2K cores, 16 ranks/node
LS3DF	8.1	ZnOTe, 8K cores, ESSLsmp, I/O sensitive

Additional BG/Q Programming Models

Model	Source	Provides	Q Status
Global Arrays Toolkit	PNNL	One sided communication and high level library for array computation	Available (MPI), PAMI pending
Berkeley UPC	Berkeley	PGAS variant of C	Ported with XLC and gcc
GASnet	Berkeley	Active message communication	Available, PAMI and MPI ports
Charm++	UIUC	Dynamic load balancing and task virtualization	Available, PAMI and MPI ports
CoArray Fortran	Rice U.	PGAS variant of Fortran	Debugging. Uses GASnet

Mira Performance Tools

BG/Q Tools & Libraries Project

Tool Name	Source	Provides	Mira Status
BGPM	IBM	HPC	Available
gprof	GNU/IBM	Timing (sample)	Available
TAU	Unv. Oregon	Timing (inst), MPI	Available
Rice HPCToolkit	Rice Unv.	Timing (sample), HPC (sample)	Available
mpiP	LLNL	MPI	Available
PAPI	UTK	HPC API	Available
Darshan	ANL	IO	Available
Open|Speedshop	Krell	Timing (sample), HCP, MPI, IO	Available
Scalasca	Juelich	Timing (inst), MPI	Available
Jumpshot	ANL	MPI	Available
DynInst	UMD/Wisc/IBM	Binary rewriter	In development
ValGrind	ValGrind/IBM	Memory & Thread Error Check	In development

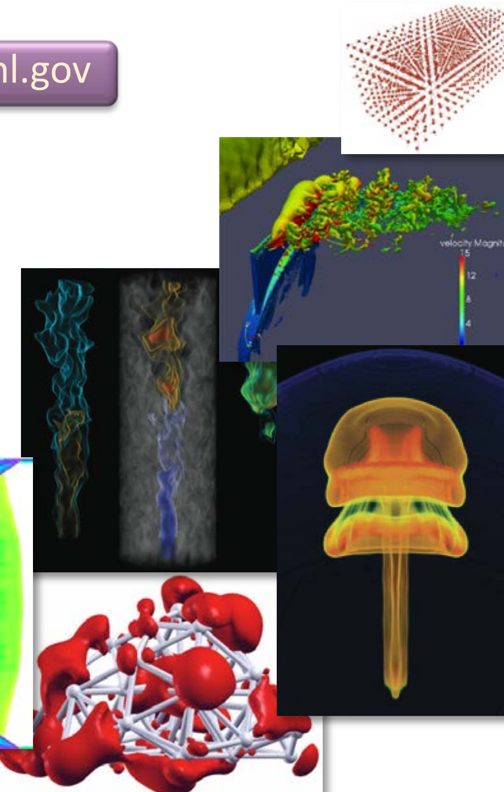
Mira Libraries

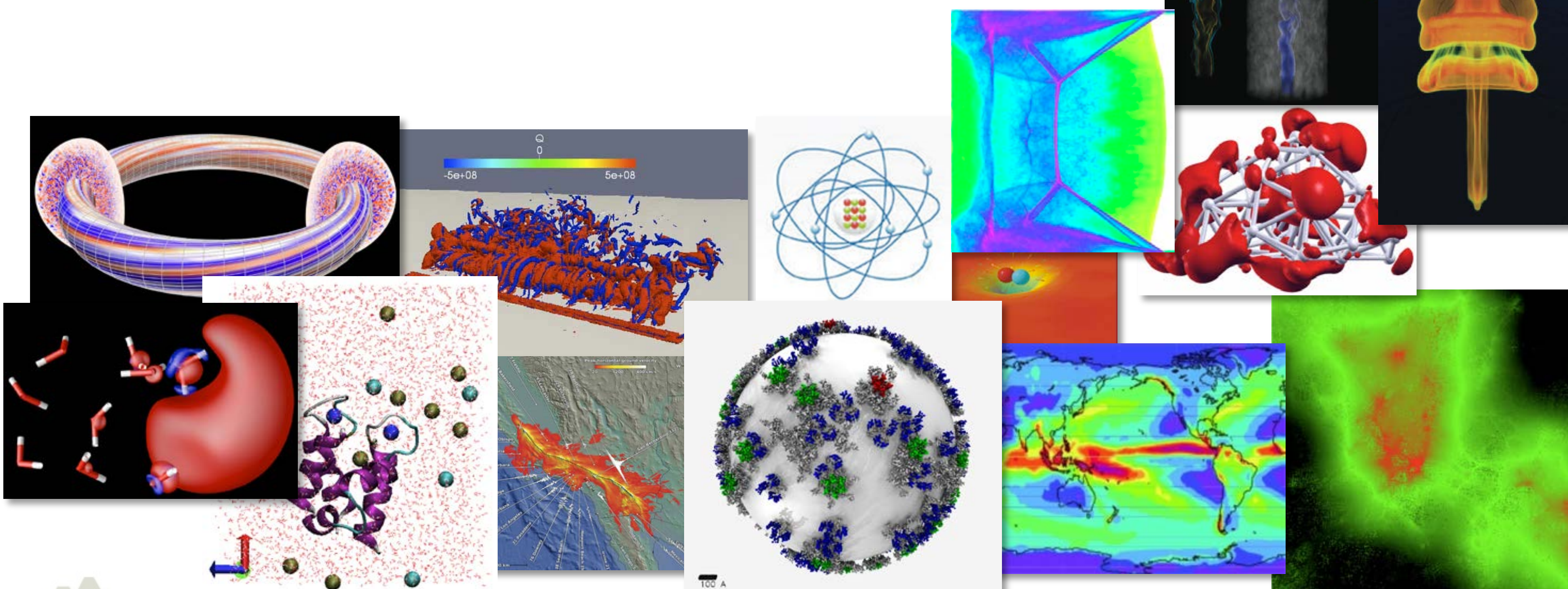
Library	Source	Provides	Mira Status
ESSL	IBM	Dense Linear Algebra & FFT Kernels	5.1.1-0 beta version
BLAS	NETLIB (UTK) & ESSL	Dense linear algebra kernels	Available, based on ESSL GEMM
BLIS	U. Texas & ANL	Framework for GotoBLAS successor	In development
CBLAS	UTK	C wrappers to BLAS	Available
LAPACK	UTK	Dense linear algebra solver	3.4.1
SCALAPACK	UTK	Parallel dense linear algebra solver	2.0.2
ARPACK & PARPACK	Rice U.	Eigenvalues & eigenvectors	2.1
FFTW	MIT	Fast Fourier transform	2.1.5, 3.3.1, no Q hand tuning yet
METIS	UMN	Graph partitioning (for meshes)	5.0.2
ParMETIS	UMN	Graph partitioning (for meshes)	4.0.2
Boost	OSS community	C++ data structures & algorithms	1.49.0
HDF5	NCSA	High-level I/O	1.8.8
NetCDF	UCAR	High-level I/O	4.1.3
P3DFFT	SDSC	Parallel FFT	2.4 (patched by ANL)
PETSc	ANL	Parallel solvers	Available, 3.3-p1, initial PTHREADs

First in *Mira* Queue: Early Science Program

http://esp.alcf.anl.gov

- **2 billion core-hours on Mira for science runs between *Mira* acceptance and start of production**





The ALCF Early Science Program (ESP)

- **The ESP was launched to**
  - prepare key applications for the architecture and scale of Mira and
  - solidify libraries and infrastructure
- **16 Early Science projects were chosen based on computational and scientific reviews.**
- **The projects have a running start for delivery of exciting new science**
  - Postdocs were assigned to the projects
  - ALCF staff has held three workshops for the ESP projects
  - Before Mira was available, development was done through
    - Time on Intrepid
    - Some access to BG/Q systems at IBM
    - Time on Vesta, a 1-rack BG/Q Test & Development system at ALCF
    - Some access to Sequoia BG/Q system at LLNL

Profile of the ESP Projects

Science Areas
Astrophysics
Biology
CFD/Aerodynamics
Chemistry
Climate
Combustion
Cosmology
Energy
Fusion Plasma
Geophysics
Materials
Nuclear Structure

Algorithms/Methods
Structured Grids
Unstructured Grids
FFT
Dense Linear Algebra
Sparse Linear Algebra
Particles/N-Body
Monte Carlo

7 National Lab PIs

9 University PIs

Timeline for Mira availability

ESP access to Mira

- **ESP projects will have access to Mira late summer 2012**
  - **access to 1-rack BG/Q system since March**, some access to racks at IBM
- **An ESP project completed a 16-rack science run late July, and a full-machine performance run on August 3rd**
- **Currently in preparation for acceptance phase**
- **Early *science* runs likely to start second half CY2012**

Plan for 2013 INCITE time on Mira

- **Mira is committed to go live for INCITE on October 1, 2013 with 768M core-hours for allocation**
- **Start date of production INCITE time is likely to happen earlier in CY 2013**
- **Guidance for INCITE 2013**
  - Propose science based on a 3B core-hour pool, 100-300M per project
  - 2013 INCITE Allocation scenarios

All ESP Projects are running on BG/Q (1)

Project Title	PI/affiliation	Code(s)	Runs on BG/Q
Climate-Weather Modeling Studies Using a Prototype Global Cloud-System Resolving Model	Balaji/GFDL	HIRAM	✔
First Principles Calculations of Interfaces in Electrical Energy Storage Systems	Curtiss/ANL	QMCPACK	✔
Direct Numerical Simulation of Autoignition in a Jet in a Cross-Flow	Frouzakis/Swiss Fed Inst of Technology	Nek5000	✔
High Accuracy Predictions of the Bulk Properties of Water	Gordon/Iowa State	GAMESS	✔
Cosmic Structure Probes of the Dark Universe	Habib/ANL, UC	HACC	✔
Accurate Numerical Simulations Of Chemical Phenomena Involved in Energy Production and Storage with MADNESS and MPQC	Harrison/ORNL	MADNESS & MPQC	✔
Petascale, Adaptive CFD	Jansen/U Colorado	PHASTA	✔
Using Multi-scale Dynamic Rupture Models to Improve Ground Motion Estimates	Jordan/USC	AWP-ODC, SORD	✔

All ESP Projects are running on BG/Q (2)

Project Title	PI/affiliation	Code(s)	Runs on BG/Q
High-Speed Combustion and Detonation	Khoklov/UC	HSCD	✔
Petascale Simulations of Turbulent Nuclear Combustion	Lamb/UC	FLASH/RTFlame	✔
Lattice Quantum Chromodynamics	Mackenzie/Fermilab	MILC, Chroma, CPS	✔
Petascale Direct Numerical Simulations of Turbulent Channel Flow	Moser/UTexas	PSDNS	✔
Ab-initio Reaction Calculations for Carbon-12	Pieper/ANL	GFMC	✔
NAMD - The Engine for Large-Scale Classical MD Simulations of Biomolecular Systems Based on a Polarizable Force Field	Roux/UC	NAMD, Charm++	✔
Global Simulation of Plasma Microturbulence at the Petascale & Beyond	Tang/Princeton	GTC-P	✔
Multiscale Molecular Simulations at the Petascale	Voth/UC, ANL	NAMD, LAMMPS, Raptor	✔

All ESP applications are running on BG/Q --

How much effort did it take?

- In the next two slides we use S = small, M = medium, L = Large as measures, characterized as
  - S = zero - few days of effort, modifications to 0 - 3% of existing lines of code
  - M = a few weeks of effort, modifications to 3% - 10% of existing lines of code
  - L = a few months of effort, changes to beyond 10% of existing lines of code

- The ranking is based on estimates by the people who actually did the work

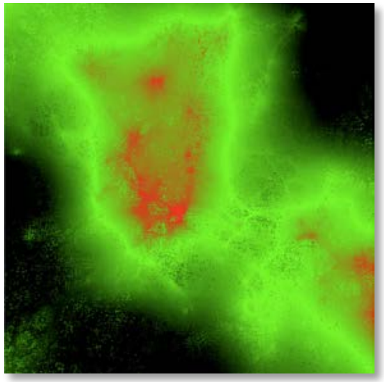
All ESP Projects are running on BG/Q (1)

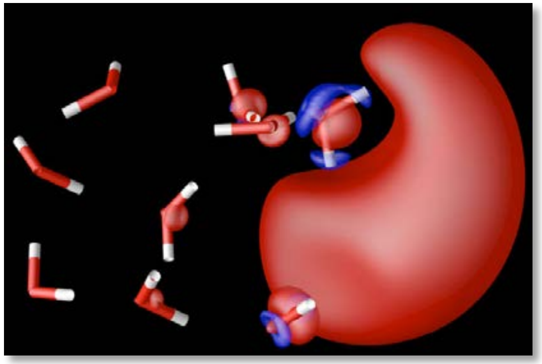
PI/affiliation	Code(s)	Runs on BG/Q	Magnitude of changes Small, Medium, Large	Nature of Changes
Balaji/GFDL	HIRAM	✔	L	Improve OpenMP implementation, reformulate divergence-damping
Curtiss/ANL	QMCPACK	✔	M	S to port, L to use QPX in key kernels; plan: nested OpenMP
Frouzakis/Swiss Fed Inst of Technology	Nek5000	✔	S	Optimized small matrix-matrix multiply using QPX
Gordon/Iowa State	GAMESS	✔	M	64-bit addressing, thread integral kernels with OpenMP
Habib/ANL, UC	HACC	✔	M	Short-range-force only: tree code
Harrison/ORNL	MADNESS	✔	S	Threading runtime tuning Kernel tuning to use QPX
Jansen/U Colorado	PHASTA	✔	S	Unchanged MPI-only performs well; OpenMP threaded in testing
Jordan/USC	AWP-ODC, SORD	✔	S, M	None, Threading

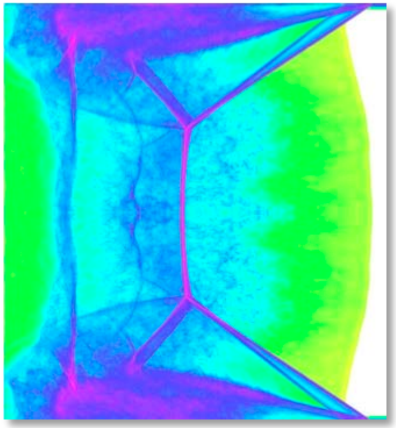
All ESP Projects are running on BG/Q (2)

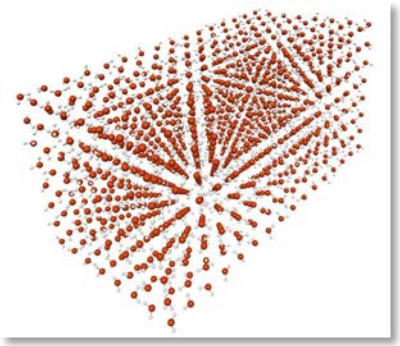
PI/affiliation	Code(s)	Runs on BG/Q	Magnitude of changes	Nature of Changes
Khoklov/UC	HSCD	✔	S	Tune OpenMP parameters, link optimized math libs
Lamb/UC	FLASH/RTFlame	✔	S	OpenMP threading
Mackenzie/Fermilab	MILC, Chroma, CPS	✔	L	Full threading, QPX intrinsics/assembler, kernel on SPI comm.
Moser/UTexas	PSDNS	✔	S	Compile dependency libs, add OpenMP directives for threading
Pieper/ANL	GFMC	✔	S	Tune no. threads & ranks.
Roux/UC	NAMD, Charm++	✔	L	Threads, PAMI implementation of Charm++
Tang/Princeton	GTC	✔	S	Improve OpenMP implementation
Voth/UC, ANL	NAMD, LAMMPS, RAPTOR	✔	M	OpenMP threads & serial optimizations in RAPTOR/LAMMPS

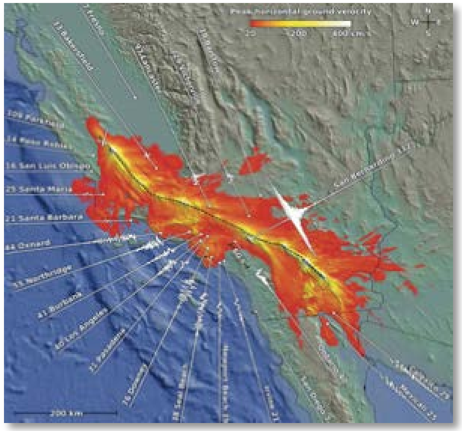
Status of selected ESP projects







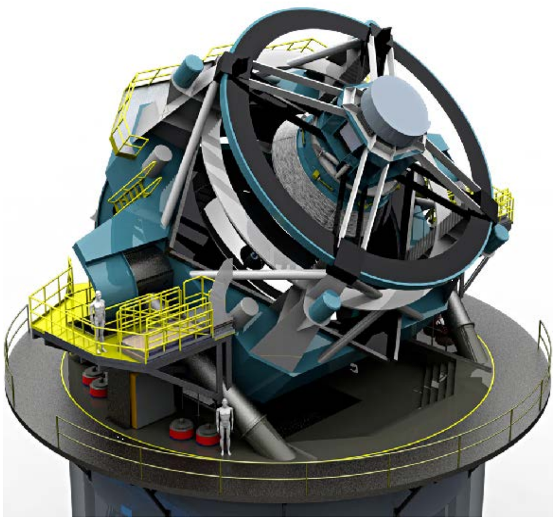


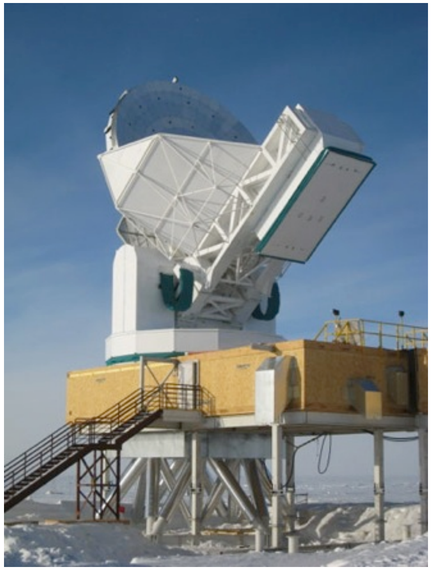


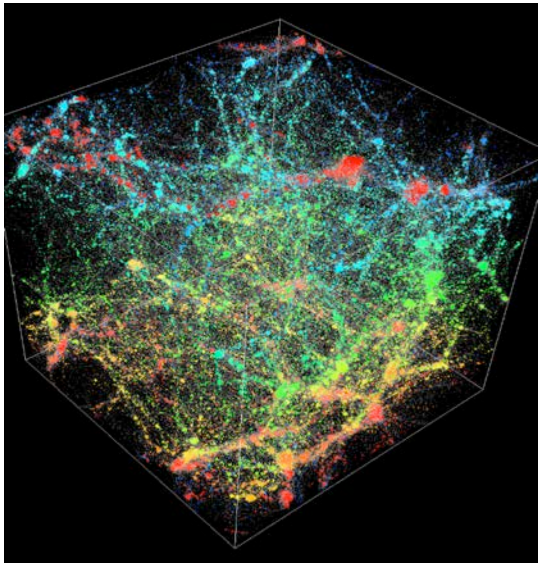
Cosmic Structure Probes of the Dark Universe

Salman Habib (Argonne National Laboratory)
Hal Finkel (ESP postdoc, ALCF)

- **Code: Hardware/Hybrid Accelerated Cosmology Code (HACC) framework**
- **Formation and evolution of large-scale structure in the Universe**
  - Characterize dark energy & dark matter by predicting observational signatures for a variety of new/existing experimental cosmological probes
    - 1^{st} simulations resolving galaxy-scale mass concentration at size scale of state-of-the-art sky surveys
    - Precision predictions for many 'sky survey' observables
  - Study primordial fluctuations by predicting the effects on cosmic structures today

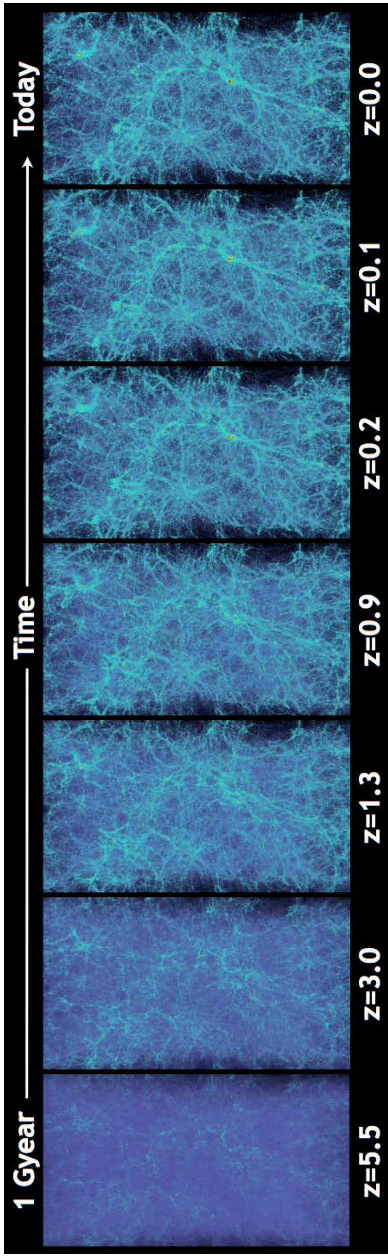






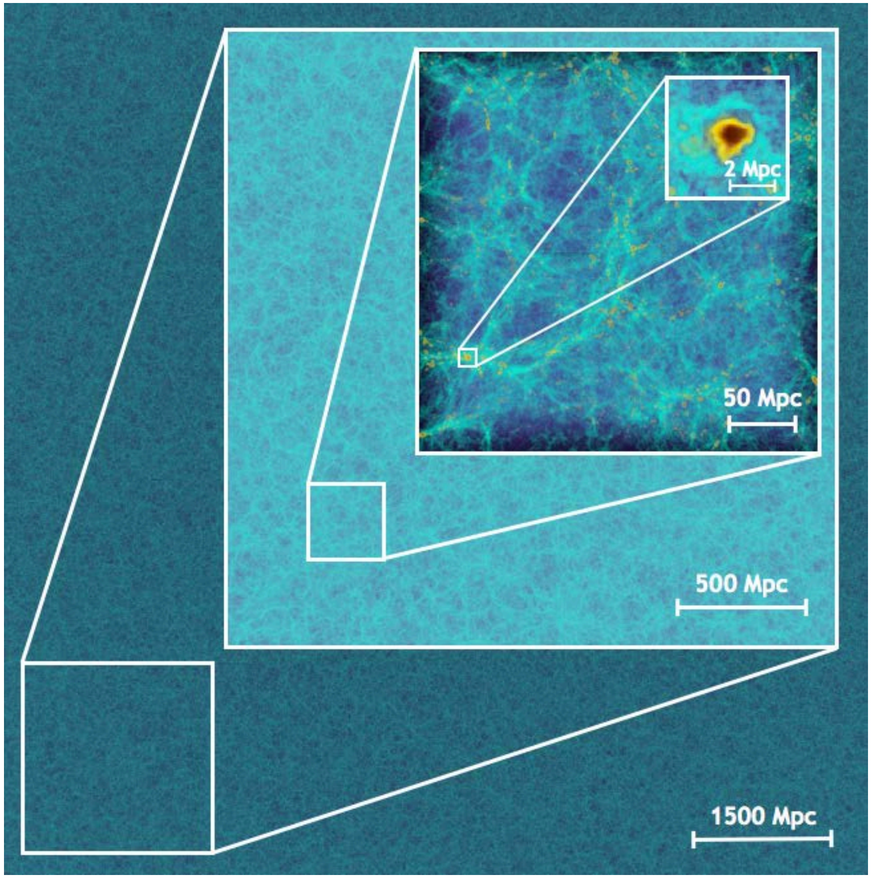
HACC on Mira

- **Hardware/Hybrid Accelerated Cosmology Code Framework**
  - Hybrid particle/grid solver for the cosmological Vlasov-Poisson system
    - Designed using practical co-design principles and associated algorithms
    - Targeted for very high performance on the largest science problems
    - Runs at scale on all current hardware (BG, CPU, CPU/GPU, CPU/MIC, --)
    - Novel tunable algorithmic structure; multiple programming paradigms
- **The Cosmological Structure Formation Problem**
  - Very high global dynamic range in space, time, and mass
    - Global spatial dynamic range requirement of 1,000,000:1
    - Mass resolution dynamic range of 100,000:1
    - Future surveys (e.g., LSST) demand multi-trillion particle simulations
- **HACC on the BG/Q**
  - Algorithms, implementations, and optimizations for the BG/Q
    - Split force computations into 3 domains, close-, medium-, and long-range
    - (i) Close -- direct Particle-Particle (PP), (ii) Medium -- Recursive Coordinate Bisection (RCB) Tree, (iii) Long -- Spectral Particle-Mesh (SPM)
    - **MPI+OpenMP model, QPX intrinsics, very high-performance force kernel**
    - **Gordon Bell Finalist: Performance at >50% of peak with >90% parallel efficiency up to 786,432 cores, multi-trillion particles at full dynamic range**



HACC on Mira: Early Science Test Run

- **First Large-Scale Simulation on Mira**
  - 16 racks, 262,144 cores (one-third of Mira)
  - 14 hours continuous, no checkpoint restarts
- **Simulation Parameters**
  - 9.14 Gpc simulation box
  - 7 kpc force resolution
  - 1.07 trillion particles
- **Science Motivations**
  - Resolve halos that host luminous red galaxies for analysis of SDSS observations
  - Cluster cosmology
  - Baryon acoustic oscillations in galaxy clustering

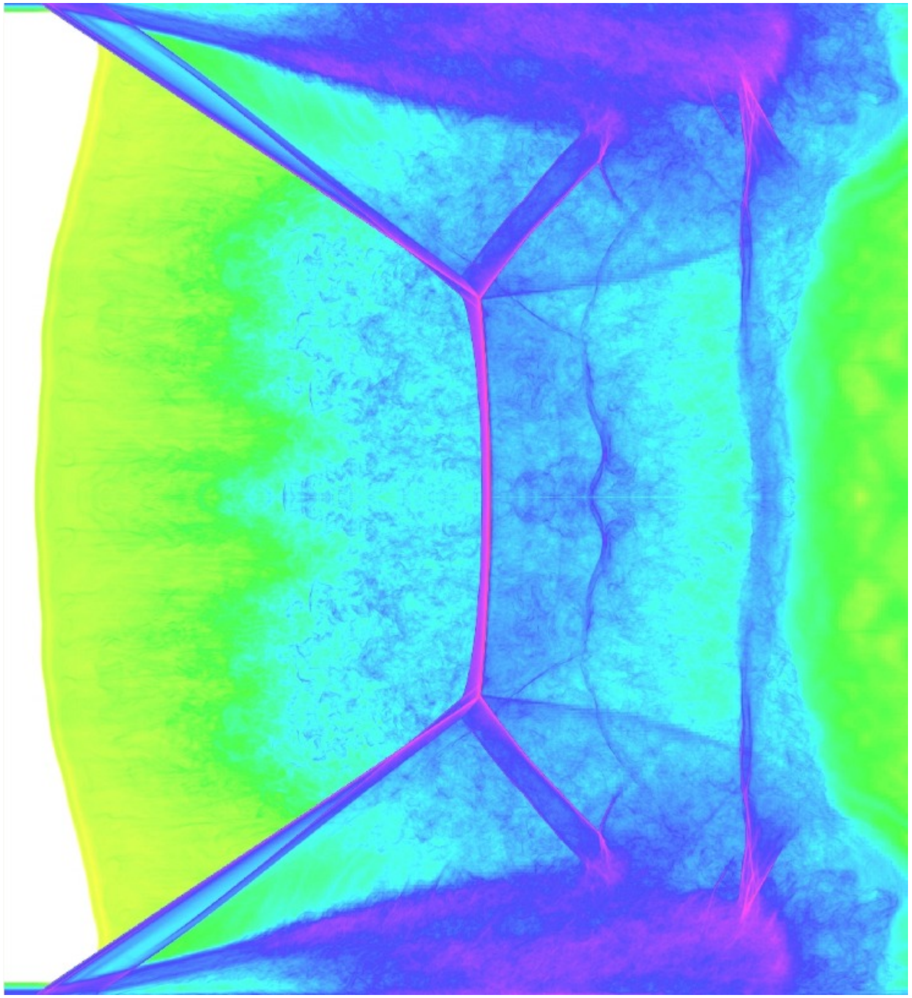


Zoom-in visualization of the density field illustrating the global spatial dynamic range of the simulation -- approximately a million-to-one

Early Science High Speed Combustion and Detonation Project

- **Direct Numerical Simulation of the deflagration-to-detonation transition (DDT) in hydrogen-oxygen gaseous mixtures for hydrogen safety**

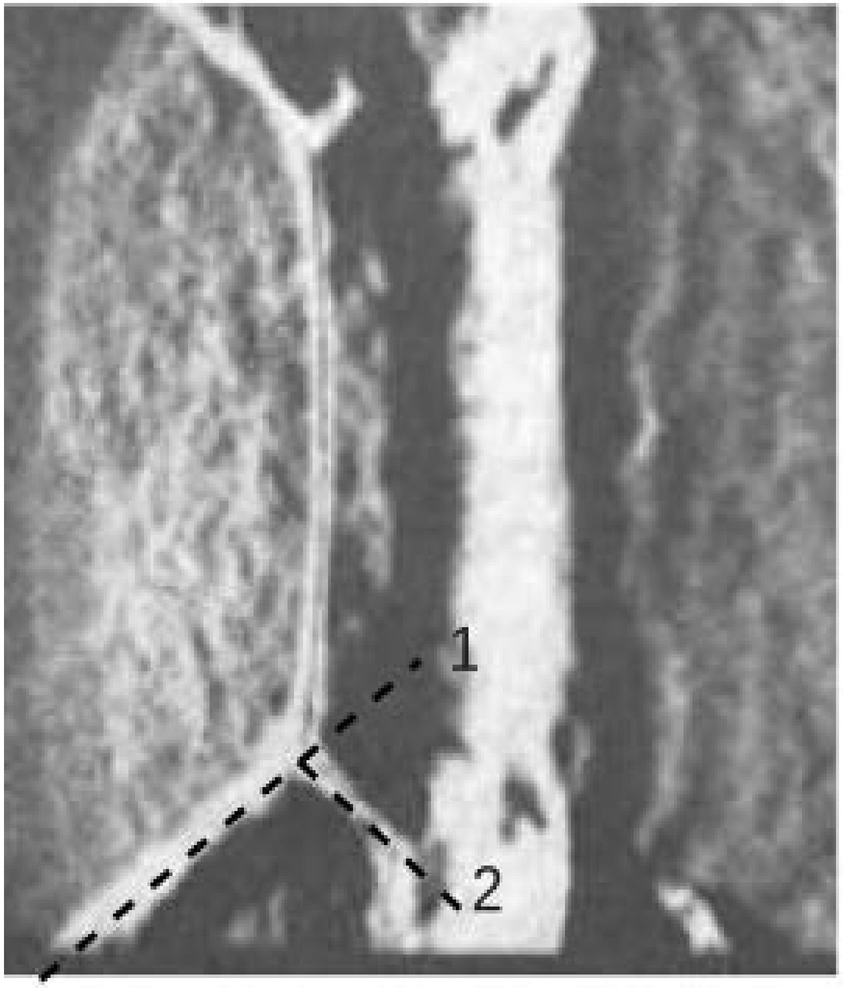
- **PI: Alexei Khokhlov, University of Chicago**
- **Joanna Austin, University of Illinois**
- **Charles Bacon, Argonne National Laboratory**
- **Andrew Knisely, University of Illinois**
- **Ben Clifford, Argonne National Laboratory**
- **Joe Bernstein, Argonne National Laboratory**

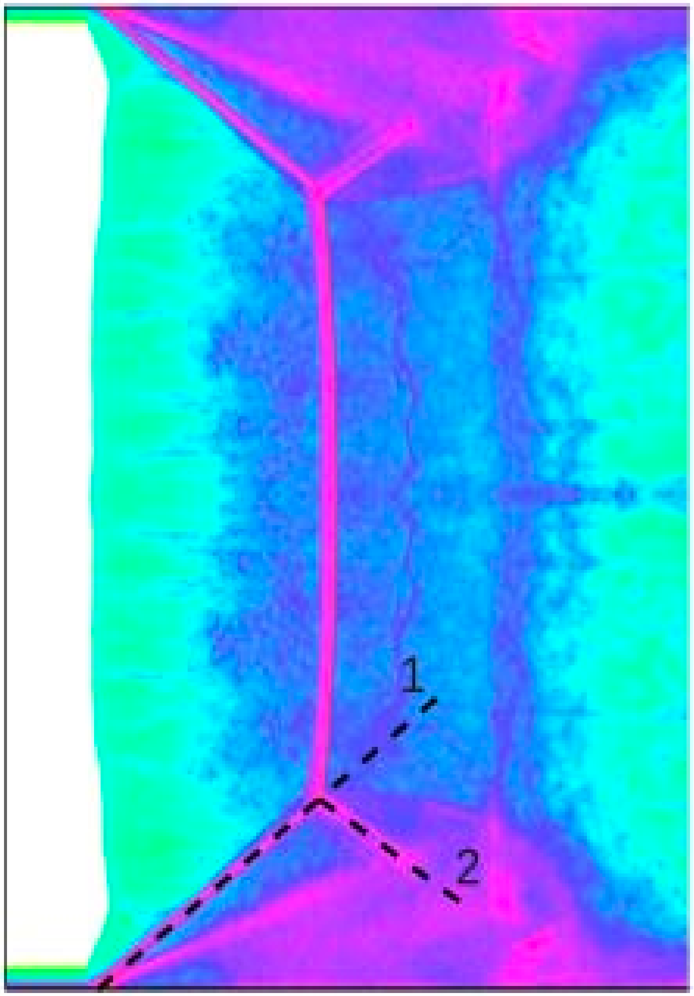


Code features

- 3-d reactive flow Navier-Stokes with 8-species and 19 reaction kinetics H2-O2 burning, multi-species NASA7 equation of state, multi-species temperature dependent viscosity, mass and heat conduction, and radiative cooling
- Adaptive mesh refinement on a regular rectangular grid

Reflected shock tube validation





Single-node scaling on Q

Thread count	Time per step	Efficiency (%)
1	349	1
2	190	92
4	110	79
8	70 (68)	62 (64)
16	64 (48)	34 (45)
32	45 (40)	24 (27)
64	50 (41)	11 (13)

Parenthetical numbers come after increasing the size of the array of cells passed to the work functions – high rank counts were getting not enough work per thread from the original setting

Main loop times

- **BG/P -> BG/Q speedup = 2.5x/core, 9.2x/node**

BG/Q Node count	Time	Efficiency
128	705	100
256	358	98.5
512	188	93.8
1024	101	87.3
BG/P Node count	**Time**	**Efficiency**
512	1779	100
1024	887	100
2048	488	91.1
4096	250	89

FLASH Overview

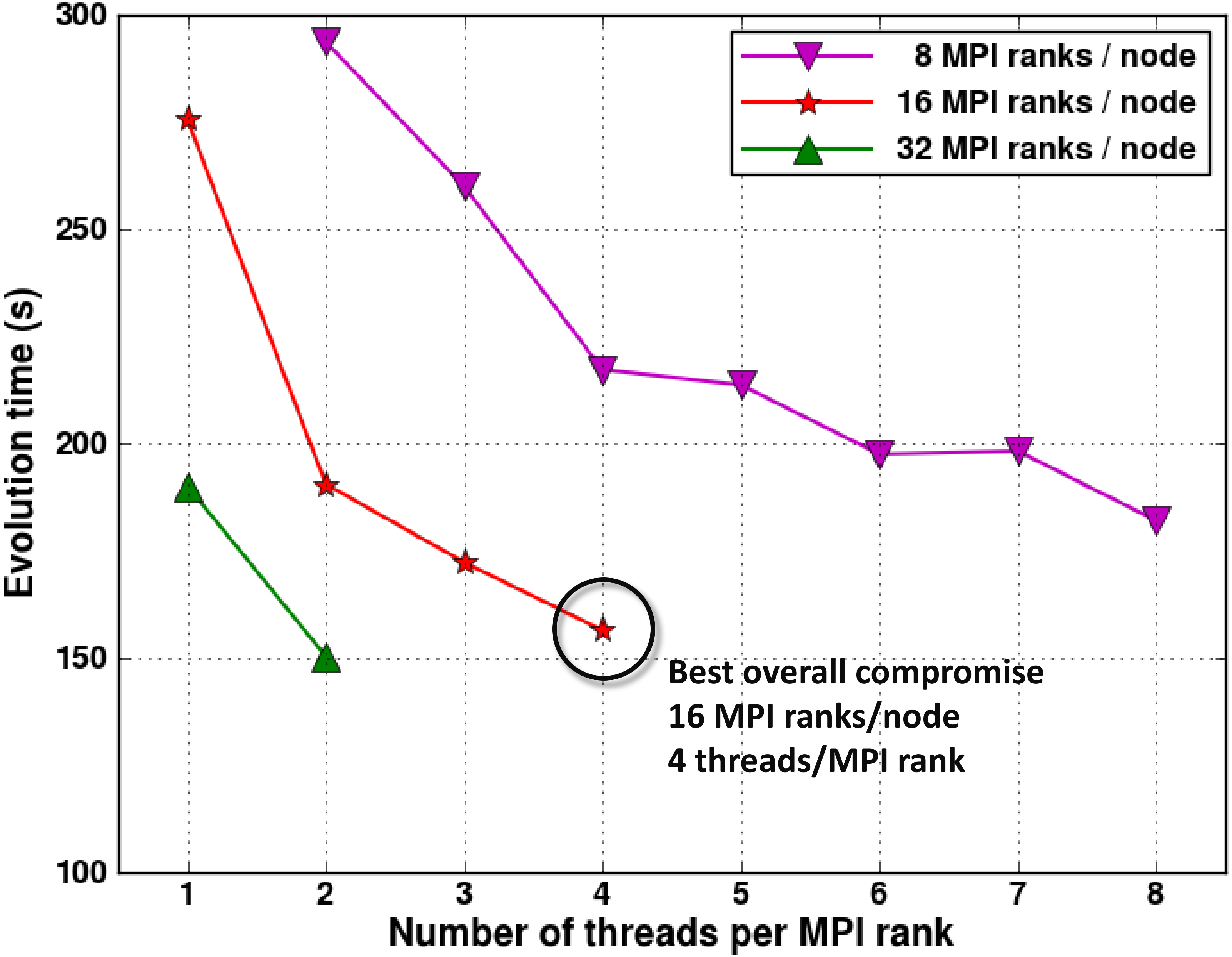
PI: Don Lamb, University of Chicago
Postdoc: Christopher Daley

- FLASH simulates problems from astrophysics, cosmology, HEDP, incompressible fluid dynamics
- Collection of code units that a user assembles into a custom application
- Portable and scalable
- ½ million lines of code
  - Fortran90 and C
  - MPI and (recently) OpenMP
  - AMR
  - Parallel I/O

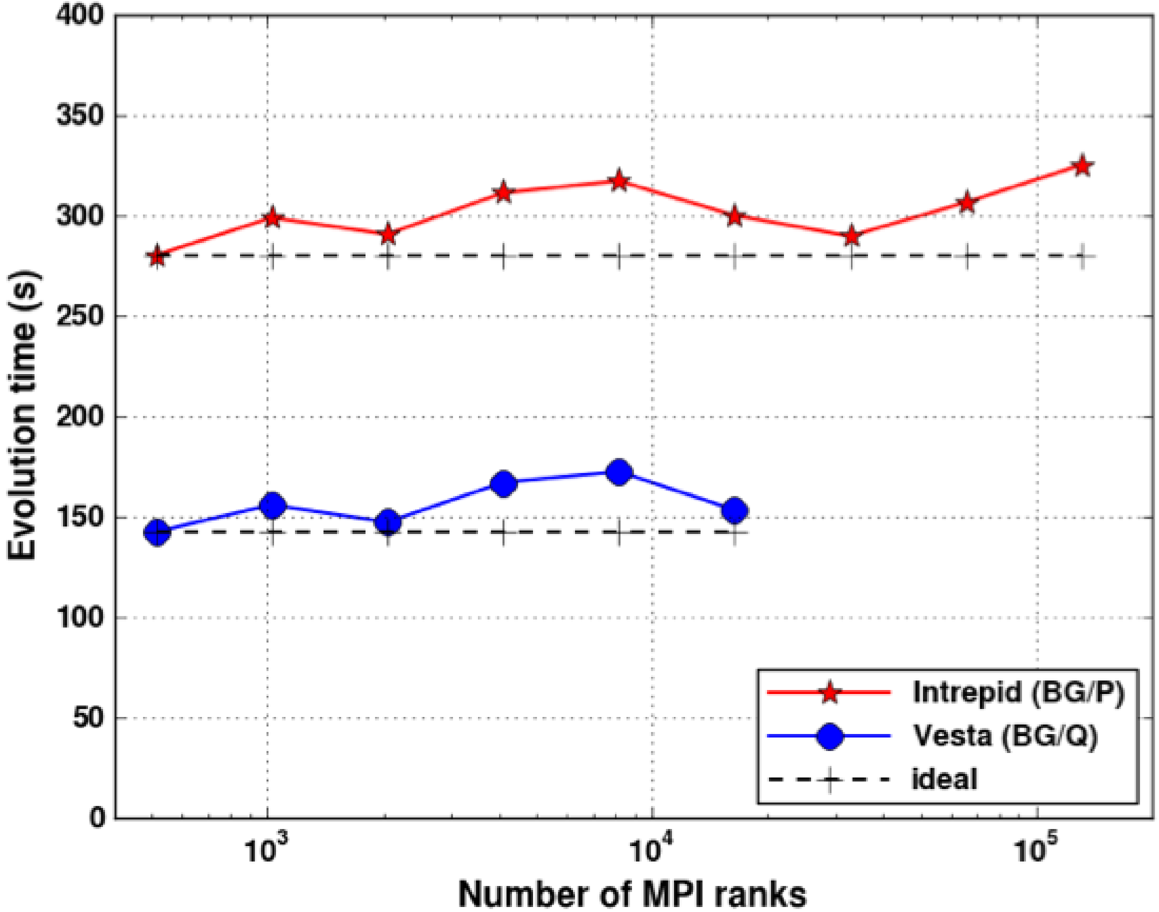
FLASH Early Science Goals

- Improve understanding of type 1a supernova
- The early science applications include
  - RTFlame (flame in a rectilinear domain with constant gravity)
  - White Dwarf (full supernova problem)

RTFlame: FLASH performance on BG/Q



White Dwarf: Weak scaling on BG/Q



Source: Chris Daley, postdoc for FLASH ESP

White Dwarf: Weak scaling on BG/Q summary

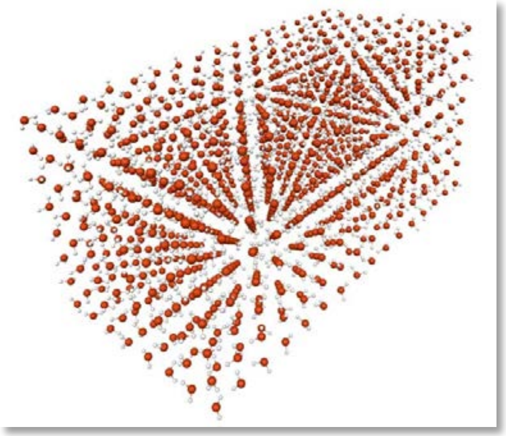
- Good weak scaling on Intrepid BG/P (up to 131K MPI ranks) and Vesta BG/Q (up to 16K MPI ranks - whole machine)
  - Hard to create perfect weak scaling adaptive grid tests
    - Oscillations in evolution time happen because the work per MPI rank does not stay exactly constant
- BG/Q to BG/P node-to-node ratio between 7.4x and 7.9x
  - Less than RTFlame node-to-node ratio
    - White Dwarf application includes additional physics and so more guard cell fills per time step
    - Does not include multipole solver multithreading - removed to avoid a strange crash during access of threadprivate data

Source: Chris Daley, postdoc for FLASH ESP

GAMESS

(General Atomic and Molecular Electronic Structure System)

- *Ab initio* quantum chemistry package
- Maintained by the research group of Prof. Mark Gordon at Iowa State
  - ESP Postdoc: Maricris Lodriguito Mayes
  - ALCF catalyst: Graham D. Fletcher
- Implements all major quantum mechanical methods
  - Hartree- Fock
  - Møller Plesset perturbation theory
  - Coupled-cluster
  - Multiconfiguration consistent field
  - Configuration interaction
  - Density functional theory



- Ported to all major architectures
- Free and widely used on everything from laptops to supercomputers
- About a million lines of code, with an associated parallelization library comprising 15,000 lines
- Highly scalable, including many distributed data algorithms

41

GAMESS BG/Q Early Access Results

The Performance of FMO2-MP2/aug-cc-pvdz force calculations for several water clusters. (All timings are in minutes)

				Number of Nodes			
System			Blue Gene/P	Blue Gene/Q*			
Waters	Atoms	Basis Functions	2048	128	256	512	1024
128	384	5 504	4.8	12.0	6.2	3.3	1.9
256	768	11 008	10.5	31.1	15.8	8.2	4.4
512	1536	22 016	28.9	94.2	48.8	24.9	12.9

** Using –O2 IBM XL compiler option and 4 ranks per node*

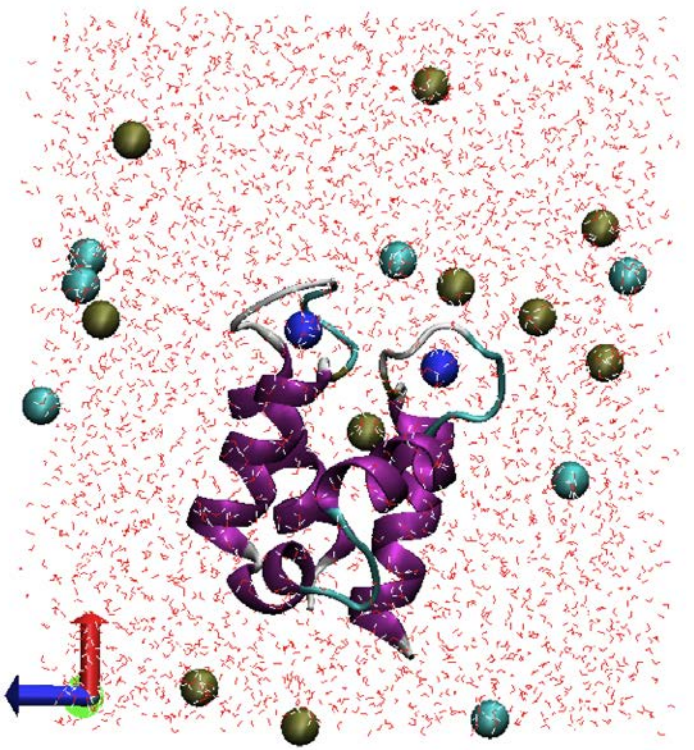
➢***Even without tuning and optimizing GAMESS, a Q/P speed-up per node is about 4.6 to 6.4 times.***

42

NAMD - The Engine for Large-Scale Classical MD Simulations of Biomolecular Systems Based on a Polarizable Force Field

Benoit Roux (U. Chicago)
Yun Luo (ESP postdoc, ALCF)
NAMD developers (Univ. of Illinois at Urbana-Champaign)

- **Next level of methods and problems incorporating new force field.**
- **Running on Blue Gene/Q**
  - Standard (non-threaded) NAMD
    - ported by researchers at IBM
- **Threading for Blue Gene/Q**
  - Threaded NAMD version developed
    - Theoretical and Computational Biophysics Group of the Beckman Institute for Advanced Science and Technology at the UIUC
  - Threaded & nonthreaded running on BG/Q
  - Underlying Charm++ framework ported
    - MPI implementation
    - PAMI implementation
    - NAMD using PAMI Charm++ 20% faster (than MPI)

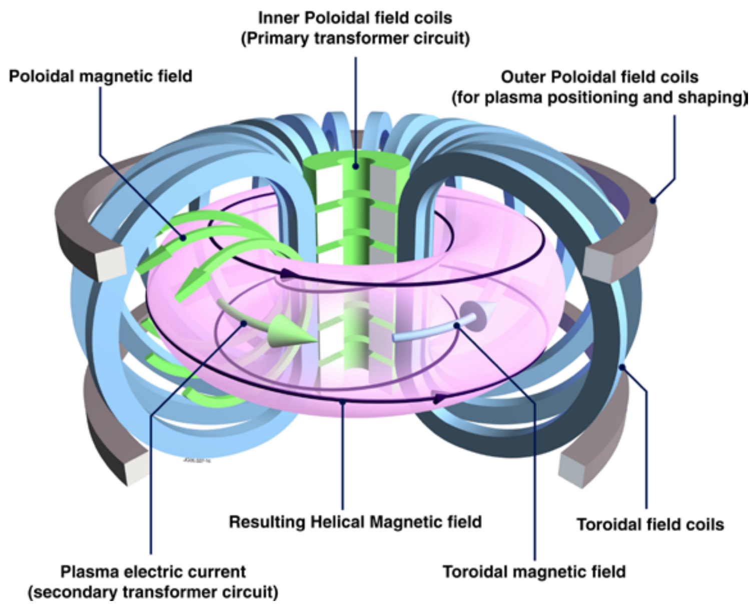


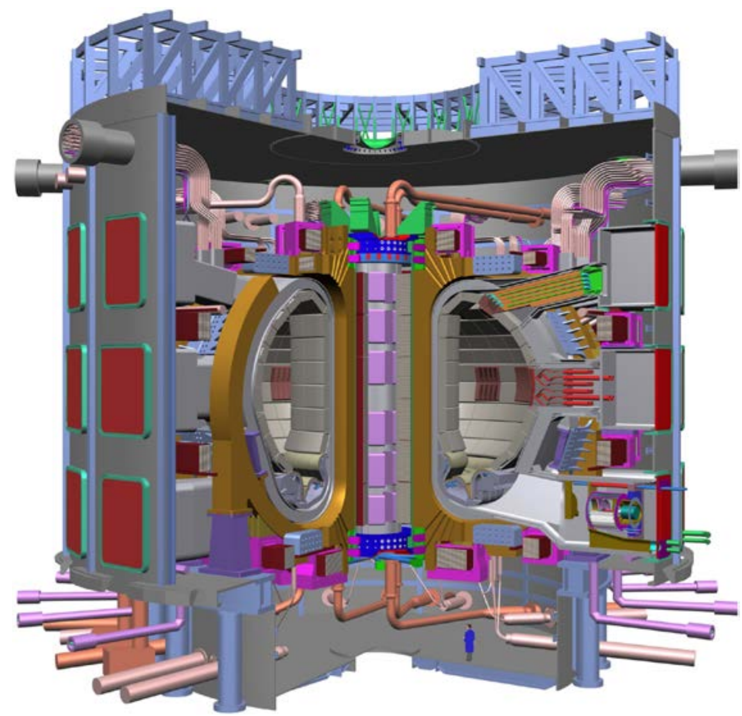
Beta Compiler Version

Tokamak Plasma Microturbulence

William Tang (Princeton U. & Princeton Plasma Physics Laboratory)
Stephane Ethier (PPPL)
Bei Wang (Princeton U.)

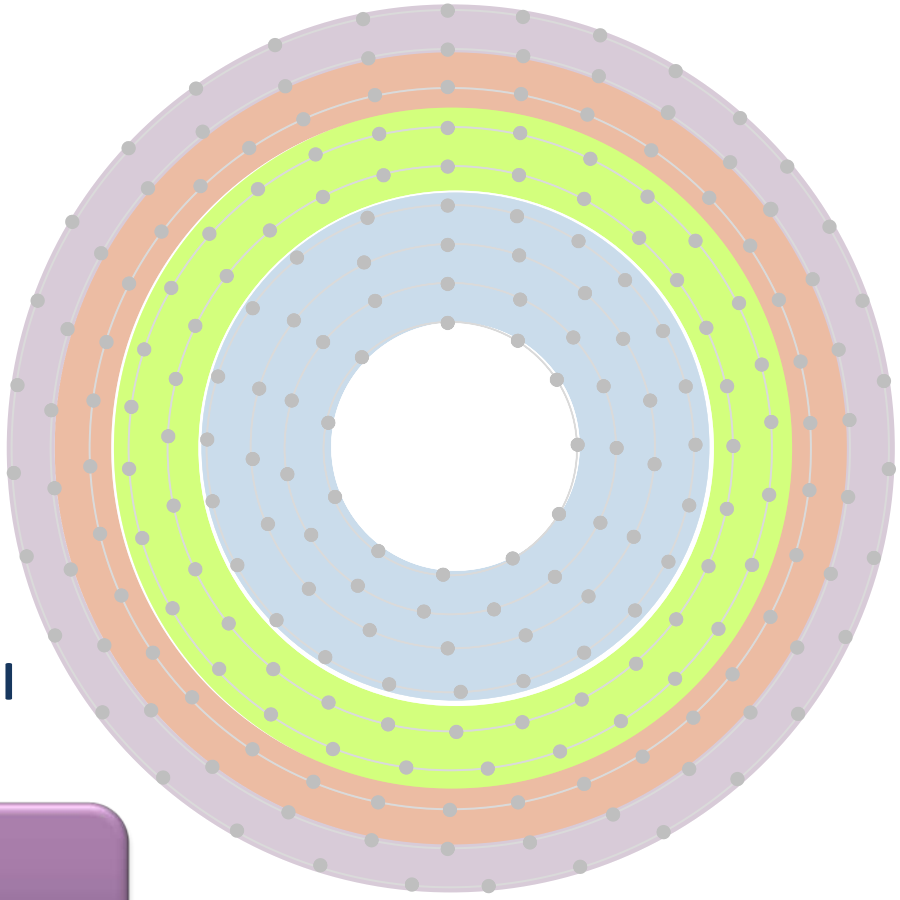
- **Code: GTC-P**
- **Particle-in-cell simulation of plasma**
  - Study energy loss through turbulence
  - Validate key assumption about scaling in ITER





Tokamak Plasma Microturbulence

- **2D data decomposition for MPI tasks**
  - Zeta (long way round torus)
    - Equal number of grid points
  - Radially in poloidal plane
    - Equal number of *particles*
    - Guard layers must overlap by max gyroradius
- **OpenMP threads**
  - Particle loops
- **~10x better performance per node (BG/Q vs BG/P)**
- **Tang: “Such reduction in time to solution will enable meaningful scientific progress.”**

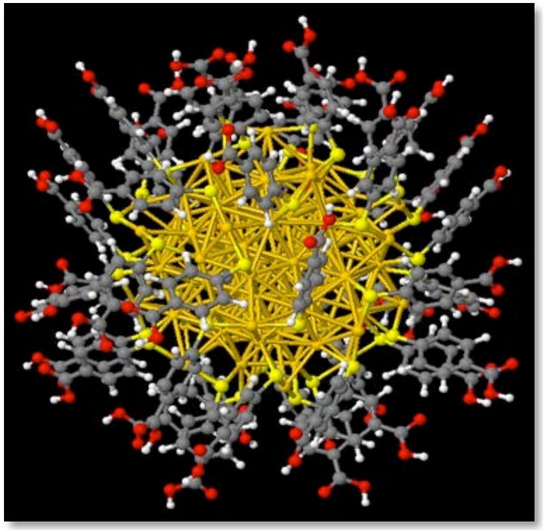


ITER plasmas
- $O(10^{10})$ particles
- $O(10^{8})$ grid cells

Materials Design and Discovery: Catalysis and Energy Storage

Larry Curtiss (Argonne National Laboratory)
Nick Romero (ALCF)
Anouar Benali (ESP postdoc, ALCF)

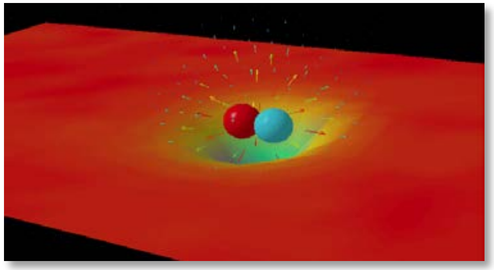
- **Electronic Structure Codes: QMCPACK, CPMD**
  - Quantum Monte Carlo (QMC)
  - Density functional theory (DFT)
- **QMCPACK on Blue Gene/Q**
  - Operations depend on type of wave function: LCAO, real-space, PWs.
    - Spline interpolation
    - Small DGEMM and DGEMV
  - Reformulating loops to use BLAS2+3
  - Adding QPX intrinsics
  - Add nested parallelism to MCWalker evaluation *(needs OpenMP ≥ 3.0)*
    - Good efficiency going from 1 to 2 OpenMP threads
    - Going from 2 to 4 threads: 1.5x speedup
- **IBM Zurich optimizing CPMD for Blue Gene/Q**



Beta Compiler Version

Lattice Quantum Chromodynamics

Paul Mackenzie (Fermilab) and the USQCD consortium
James Osborn (ALCF)
Heechang Na (ESP postdoc, ALCF)
4-years with IBM on BG/Q: {Brookhaven, Columbia U., U. Edinburgh}



- **Code: MILC, Chroma, CPS**
- **4D space-time lattice solving theory of quarks and gluons**
- **Node-level optimizations for Blue Gene/Q**
  - SIMD optimizations – quad FPU on BG/Q
  - Designed prefetching interface between CPU and L2 cache on BG/Q (Boyle/Christ/Kim)
  - Important feedback from LQCD performance to memory system design
  - Full hybrid Monte Carlo evolution now running on BG/Q
    - Dirac matrix solver kernel:
      - 60% of theoretical peak on BG/Q chip
      - >80% of theoretical peak communication bandwidth BG/Q chip
        - » Using low level SPI communications layer
  - Dirac solver kernel using pthreads
  - OpenMP used in rest of code

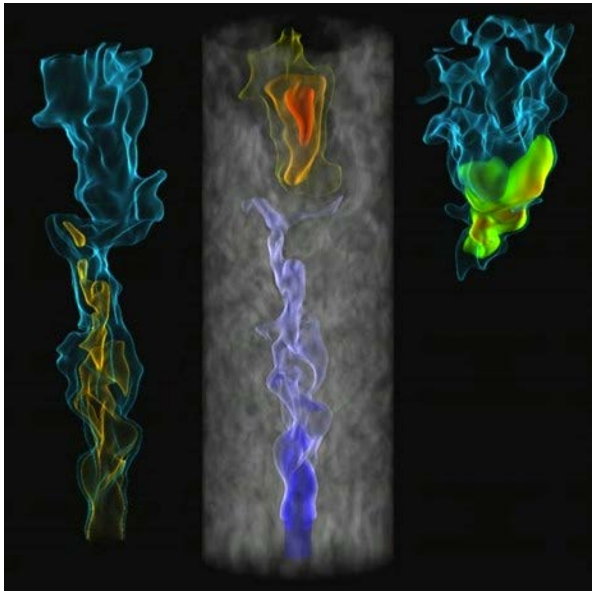
Beta Compiler Version

Direct Numerical Simulation of Autoignition in a Jet in a Cross-Flow

Christos Frouzakis(ETH Zürich)
Paul Fischer (Argonne National Laboratory)
Scott Parker (ALCF)

- **Auto-ignition of fuel-air mix related to lean combustion gas turbines**
  - Goal: avoid autoignition for safer, cleaner lean combustion
  - First simulation in lab-scale jet
  - Never studied, experimentally or computationally
- **Code: Nek5000**
  - Spectral element
- **Running on Blue Gene/Q**
  - MPI everywhere
  - Good performance up to 4 MPI ranks/core (*64 ranks/node*)
  - 6.7x better performance/node than Blue Gene/P
    - 1024 cores
    - 32 ranks/node
    - *No QPX yet*

O(10^{10}) gridpoints
100k timesteps



Beta Compiler Version

Timeline for Mira availability

ESP access to Mira

- **ESP projects will have access to Mira late summer 2012**
  - access to 1-rack BG/Q system since March, some access to racks at IBM
- **An ESP project completed full-machine performance run on DATE XX, 16-rack science run on DATE YY**
- **Currently in preparation for acceptance phase**
- **Early *science* runs likely to start second half CY2012**

Plan for 2013 INCITE time on Mira

- **Mira is committed to go live for INCITE on October 1, 2013 with 768M core-hours for allocation**
- **Start date of production INCITE time is likely to happen earlier in CY 2013**
- **Guidance for INCITE 2013**
  - Propose science based on a 3B core-hour pool, 100-300M per project
  - 2013 INCITE Allocation scenarios

Summary

- Early experience confirms that Mira will enable advances in a broad spectrum of applications
- The Early Science Program is paying off
  - All the applications are running
  - Many valuable insights on tuning and scaling are being obtained
  - Applications are being enhanced to model more complex phenomena, with higher fidelity
- The BG/Q Tools and Libraries Project has yielded substantial software tools very early in the life of Mira
- We look forward to exciting scientific results in the next few months