Blue Gene Science Applications

Dr. George Chiu
IEEE Fellow
IBM T.J. Watson Research Center
Yorktown Heights, NY

Nuclear Energy
Fluid Dynamics
Materials Science
Biological Modeling

Renewable Energy
Climate Modeling
Geophysical Data Processing
Financial Modeling
Science & Technology

- Costs of experiment and theory remain high.
- Computer simulation is the only leg of science that reduces cost according to Moore's law => Any aspect that can be simulated, should be simulated.
- If you can simulate, you should.
### IBM System Technology Group

**Examples of Applications Running on Blue Gene**

Developed on L, P; many ported to Q

<table>
<thead>
<tr>
<th>Application</th>
<th>Owner</th>
<th>Application</th>
<th>Owner</th>
<th>Application</th>
<th>Owner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biology: LifeV</td>
<td>CSCS</td>
<td>DFT BigDFT</td>
<td>CSCS</td>
<td>BM: SPEC2006, SPEC openmp</td>
<td>SPEC</td>
</tr>
<tr>
<td>CFD Alya System</td>
<td>Barcelona SC</td>
<td>DFT CP2K</td>
<td>Jülich</td>
<td>BM: NAS Parallel Benchmarks</td>
<td>NASA</td>
</tr>
<tr>
<td>CFD (Flame) AVBP</td>
<td>CERFACS Consortium</td>
<td>DFT GPAW, ls3df</td>
<td>Argonne National Lab</td>
<td>BM: RZG (AIMS,Gadget,GENE, GROMACS,NEMORB,Octopus, Vertex)</td>
<td>RZG</td>
</tr>
<tr>
<td>CFD dns3D</td>
<td>Argonne National Lab</td>
<td>DFT PARATEC</td>
<td>NERSC / LBL</td>
<td>Coulomb Solver - PEPC</td>
<td>Jülich</td>
</tr>
<tr>
<td>CFD OpenFOAM</td>
<td>SGI</td>
<td>DFT CPMD</td>
<td>IBM/Max Planck</td>
<td>MPI PALLAS</td>
<td>UCB</td>
</tr>
<tr>
<td>CFD NEK5000, NEKTAR</td>
<td>Argonne, Brown U</td>
<td>DFT QBOX</td>
<td>LLNL</td>
<td>Mesh AMR</td>
<td>CCSE, LBL</td>
</tr>
<tr>
<td>CFD OVERFLOW</td>
<td>NASA, Boeing</td>
<td>DFT VASP</td>
<td>U Vienna &amp; Duisburg</td>
<td>PETSC</td>
<td>Argonne National Lab</td>
</tr>
<tr>
<td>CFD Saturne</td>
<td>EDF</td>
<td>Q Chem GAMESS</td>
<td>Ames Lab/Iowa State</td>
<td>MpiBlast-pio Biology</td>
<td>VaTech / ANL</td>
</tr>
<tr>
<td>CFD LBM</td>
<td>Erlanger-Nuremberg</td>
<td>Nuclear Physics GFMC</td>
<td>Argonne National Lab</td>
<td>RTM – Seismic Imaging</td>
<td>ENI</td>
</tr>
<tr>
<td>MD Amber</td>
<td>UCSF</td>
<td>Neutronics SWEEP3D</td>
<td>LANL</td>
<td>Supernova la FLASH</td>
<td>Argonne National Lab</td>
</tr>
<tr>
<td>MD Dalton</td>
<td>Univ Oslo/Argonne</td>
<td>QCD CPS</td>
<td>Columbia U/IBM</td>
<td>Ocean HYCOM</td>
<td>NOPP / Consortium</td>
</tr>
<tr>
<td>MD ddcMD</td>
<td>LLNL</td>
<td>QCD MILC</td>
<td>Indiana University</td>
<td>Ocean POP</td>
<td>LANL/ANL/NCAR</td>
</tr>
<tr>
<td>MD LAMMPS</td>
<td>Sandia National Labs</td>
<td>Plasma GTC</td>
<td>PPPL</td>
<td>Weather/Climate CAM</td>
<td>NCAR</td>
</tr>
<tr>
<td>MD MP2C</td>
<td>Jülich</td>
<td>Plasma GYRO (Tokamak)</td>
<td>General Atomics</td>
<td>Weather/Climate Held-Suarez Test</td>
<td>GFDL</td>
</tr>
<tr>
<td>MD NAMD</td>
<td>UIUC/NCSA</td>
<td>KAUST Stencil Code Gen</td>
<td>KAUST</td>
<td>Climate HOMME</td>
<td>NCAR</td>
</tr>
<tr>
<td>MD Rosetta</td>
<td>U Washington</td>
<td>BM:sppm,raptor,AMG,IRS,sphot</td>
<td>Livermore</td>
<td>Weather/Climate WRF, CM1</td>
<td>NCAR, NCSA</td>
</tr>
</tbody>
</table>

**Accelerating Discovery and Innovation in:**

- **Materials Science**
- **Energy**
- **Engineering**
- **Climate & Environment**
- **Life Sciences**

- **Silicon Design**
- **Next Gen Nuclear**
- **High Efficiency Engines**
- **Oil Exploration**
- **Whole Organ Simulation**
### Overall Efficiencies of BG Applications - Major Scientific Advances

<table>
<thead>
<tr>
<th>Application</th>
<th>Efficiency</th>
<th>Award Year</th>
<th>Scaling</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Qbox (DFT) LLNL</strong></td>
<td>56.5%</td>
<td>2006</td>
<td>128Ki L, 64Ki P</td>
</tr>
<tr>
<td>CPMD IBM</td>
<td>30%</td>
<td></td>
<td>128Ki L</td>
</tr>
<tr>
<td>KKRnano</td>
<td></td>
<td></td>
<td>288Ki P</td>
</tr>
<tr>
<td><strong>ddcMD (Classical MD) LLNL</strong></td>
<td>27.6%</td>
<td>2005</td>
<td>128Ki L</td>
</tr>
<tr>
<td>New ddcMD LLNL</td>
<td>17.4%</td>
<td>2007</td>
<td>208Ki L</td>
</tr>
<tr>
<td>LAMMPS SNL</td>
<td></td>
<td></td>
<td>256Ki Q &gt;1M MPI</td>
</tr>
<tr>
<td>MUPHY (LB/MD) for hemodynamics</td>
<td>6.4%</td>
<td></td>
<td>288Ki P</td>
</tr>
<tr>
<td>RXFF, GMD</td>
<td></td>
<td></td>
<td>128Ki L</td>
</tr>
<tr>
<td>AMBER</td>
<td></td>
<td></td>
<td>8Ki L</td>
</tr>
<tr>
<td><strong>Quantum Chromodynamics CPS</strong></td>
<td>30%</td>
<td>2006 GB</td>
<td>128Ki L,P;Q</td>
</tr>
<tr>
<td>DESY</td>
<td>20%</td>
<td></td>
<td>288Ki P</td>
</tr>
<tr>
<td><strong>sPPM (CFD) LLNL</strong></td>
<td>18%</td>
<td></td>
<td>128Ki L,P;Q</td>
</tr>
<tr>
<td>LBM, Erlangen-Nuremberg</td>
<td>4.6%</td>
<td></td>
<td>288Ki P</td>
</tr>
<tr>
<td>DNS3D</td>
<td></td>
<td></td>
<td>128Ki P;Q</td>
</tr>
<tr>
<td>NEK5 (Thermal Hydraulics) ANL</td>
<td>19.3%</td>
<td></td>
<td>256Ki P;Q</td>
</tr>
<tr>
<td><strong>ParaDîs (Dislocation Dynamics) LLNL</strong></td>
<td></td>
<td></td>
<td>128Ki L</td>
</tr>
<tr>
<td><strong>WRF (Weather) NCAR</strong></td>
<td>10%</td>
<td></td>
<td>128Ki L</td>
</tr>
<tr>
<td>POP (Oceanography)</td>
<td></td>
<td></td>
<td>16Ki L</td>
</tr>
<tr>
<td>HOMME (Climate) NCAR</td>
<td>12%</td>
<td></td>
<td>32Ki L, 96Ki P</td>
</tr>
<tr>
<td><strong>GTC (Plasma Physics) PPPL</strong></td>
<td>7%</td>
<td></td>
<td>40Ki L, 128Ki P;Q</td>
</tr>
<tr>
<td>FLASH (Supernova Ia)</td>
<td></td>
<td></td>
<td>128Ki L, 160Ki P;Q</td>
</tr>
<tr>
<td>Cactus (General Relativity)</td>
<td></td>
<td></td>
<td>32Ki L, 128Ki P</td>
</tr>
<tr>
<td><strong>MUPHY (Plaque Formation) Harvard</strong></td>
<td>6.4%</td>
<td></td>
<td>288Ki P</td>
</tr>
<tr>
<td>Argonne v18 Nuclear Potential, GFMC</td>
<td>16%</td>
<td>2010</td>
<td>128Ki L</td>
</tr>
<tr>
<td>“Cat” Brain</td>
<td></td>
<td></td>
<td>144Ki P</td>
</tr>
<tr>
<td>FFT</td>
<td></td>
<td></td>
<td>256Ki P</td>
</tr>
</tbody>
</table>
Computational Fluid Dynamics
Turbulence Models

- **RANS (Reynold Averaged Navier Stokes)**
  - Overflow (L: 2Ki, P: 4Ki, Q) - Boeing
  - OpenFOAM (P,Q) – Open Source
  - Saturne (P: 8K, Q) - EDF
  - ISV: Fluent (Ansys), STAR-CD (CD-adapco), CFX (Ansys) – no source code

- **DES (Detached Eddy Simulation)** hybrid of RANS and LES
  - OpenFOAM (P,Q)

- **LES (Large Eddy Simulation)**
  - AVBP (L,P: 4Ki) - CERFACS
  - Miranda (L: 128Ki) - Livermore
  - NEK5 (L: 8Ki, P: 128Ki,Q) - Argonne
  - OpenFOAM (P,Q)
  - Saturne (P: 8K, Q)
  - WRF (L,P: 128K, Q) - NCAR

- **DNS (Direct Numerical Simulation)**
  - DNS3D (L: 128Ki, Q) - Argonne
  - NEK5 (L: 8Ki, P: 128Ki, Q)
  - OpenFOAM (P,Q)
Large Eddy Simulation of Two Phase Flow Combustion in Gas Turbines

Thierry Poinsot, CERFACS

- Code: AVBP – LES code
- 3D unstructured (full tetrahedral) mesh
- Improve airplane and helicopter engine designs – reduce design cost, improve efficiency
- Study the two-phase flows and combustion in gas turbine engines
- Results:
  - First and only simulations of a full annular chamber
  - Identified instability mechanisms that reduce efficiency in gas turbines
  - Scaled to 16Ki cores in 2009
AVBP Movie
Large scale vortex particle Direct Numerical Simulation of aircraft wakes

Aircraft wakes are a potential hazard for oncoming aircraft, responsible for air pollution and at the same time inherently linked with the fact that airplanes get airborne. Vortex wake structures have a long residence time imposing stringent safety distances between aircraft, thence limiting the take-off and landing capacities of airport runways. Wake alleviation schemes have focused on the wake susceptibility to instabilities.

In this work wake instabilities are studied through:

• A hybrid mesh-particle vortex method (1) built on top of the ppm library (2)
• An implementation on the massively parallel architecture IBM Blue Gene
• An effective scalability of the scheme that enables unprecedented billion particle simulations (up to 7 billions).

This tool provides an accurate reference Direct Numerical Simulation of the dynamics of wakes.

Large scale vortex particle Direct Numerical Simulation (DNS) of aircraft wakes

In this work wake instabilities are studied through:

- A hybrid mesh-particle vortex method (1) built on top of the ppm library(2)
- An implementation on the massively parallel architecture IBM Blue Gene
- An effective scalability of the scheme that enables unprecedented billion particle simulations (up to 7 billions 16K nodes).

Philippe Chatelain*, Alessandro Curioni**, Michael Bergdorf*, Diego Rossinelli*, Wanda Andreoni**, Petros Koumoutsakos*

*Computational Science - ETH Zurich, Switzerland
**Computational Sciences - IBM Research Zurich, Switzerland
Blue Gene, thanks to its very good scalability properties, allowed us to perform the largest ever DNS simulation for a vortex method (7 Billions of particles on 16K nodes). Thanks to these high resolution-high accuracy vortex DNS simulations, we can access details of the mechanism behind the destruction of trailing vortices, that we would have missed completely otherwise.
NASA Overflow 2.0aa Code

- **Code description**
  - Multi-block body-fitted structured overset grid code
  - Finite-volume ADI approximate factorization scheme
  - Linear step – scalar penta-diagonal solver
  - Applications (external aero, subsonic transport, rotorcraft, space shuttle, inlet)

- **Computational Details**
  - TI-08 large problem
  - 127 million grid points, 39 Grid blocks
  - 1500 time steps
  - Minor modifications to original code

- **Performance**
  - Time per time step
  - Very good scaling to 4096 cores
  - Blue Gene/P in the SMP mode
  - Blue Gene/L in the Virtual Node Mode
  - Exploits dual modes of parallelism
Climate, Weather, Oceanography
HOMME and CCSM on Argonne BG/P

- **Aqua planet experiment runs**
  - Full physics, no land model

- **Excellent performance to scale between L and P**
  - P runs in VN mode (56K cores)
  - L runs in CO mode (98K cores)

- **BG/P, at 0.50 degree, achieves an integration rate of over 20 Simulated Year Per Day**
Weather Modeling - WRF

WRF: Weather Research and Forecasting Model
Used for both research and operational forecasting
Scaled to 128K cores on BG/P
POP: Parallel Ocean Program (NCAR)

- Simulates Global Eddies in World’s Oceans.
- Major component in climate models
- Goal is to “hindcast” 55 year ocean behavior from 1955 – present.
- Originally run at LLNL, but showed poor performance.
- After development effort from NCAR, scaling problems resolved, and tested during BGW day.
- On 16 Racks, 55 year hindcast can be completed in 8 days!

- One short term possibility – RFP for Tsunami Early Warning Systems. RFP Discussions beginning
  - Dr. J. Sairamesh (Ramesh),
Lattice Boltzmann Method –
Fully Resolved Rigid Bodies Immersed in Fluid Flow
Jan Götz, Klaus Iglberger, Marus Stürmer, Ulrich Rüde –
University of Erlangen-Nuremberg, Germany

- Weak scaling to full 72 racks (294,912 cores) of BG/P at Jülich with 0.264 billion particles.
- Fully resolved rigid bodies (particles).
- Explicit coupling algorithm that transfer momenta from the fluid to particles in each time step.
- Particles impose moving boundaries for the fluid solver.
- Red particles lighter than water (0.8 gm/cm³), float to the top.
- Brown rods heavier than water (1.2 gm/cm³), sink to the bottom.
Fluid Flow and Mass Transport in Porous Media –
Siarhei Khirevich, Anton Daneyko –
Philipps University of Marburg, Germany

- Important for many industrial processes such as: chromatographic separations, reaction engineering, catalysis, soil remediation, etc.
- Strong scaling to full 72 racks (294,912 cores) of BG/P at Jülich
- Lattice Boltmann Method (LBM) for fluid flow
- Random Walk Particle Tracking (RWPT) tracking method for advective-diffusive mass transport
- MPI I/O is used to achieved 9GB/s, 30% of the total file system I/O bandwidth.

Figure 1: a) Performance scaling on JUGENE system. The upper number in each pair indicates the ratio of the longitudinal dimensions of the longest to the shortest decomposed subdomains. The lower number is the fraction of the processes with the longest domain length. b) I/O performance of LBM (file write) and RWPT (file read).
Thermal Hydraulics

- CFD + heat transfer
Current frontier with *Code_Saturne*

Calculation under way with 100 million cells

- PWR assembly mixing grid
- calculation on 4 000 to 8 000 procs
- major lock due to mesh generation
Mixing grid meshing

- 100 million tetrahedral meshing, edge length: 0.2 mm
- 8000 cores Saturne calculation (k-ε), 15000 iterations to reach convergence
colored: vertical velocity streamlines
white: streamlines of the projected velocity on a horizontal plane
High Precision Thermal-Hydraulic Calculations Within Core
Simulation promises ...

- More accurate simulation of hot and cold water mixing in the core vessel adds 10 years to lifetime margins.

Simulation versus Design = + 10 years!
Spartacus

Physical features
- Fully Lagrangian: no mesh (SPH method)
- Weakly compressible flows
- Applicability: free-surface or confined flows, moving bodies
- Multiphase flow: only with fluids of close densities
- Turbulence modelling: state-of-the-art
- Consistency: fully conservative

Numerics and environment
- Parallelism: MPI for massive clusters
- Post-processing: Rubens, Tecplot, Paraview
- Validation: 5 test cases in 2-D, 4 in 3-D
- Documentation: 170 pages document (only for 2-D)

IBM validated up to 16Ki cores with 25M particles
Molecular Dynamics
Classical MD – ddcMD
2005 Gordon Bell Prize Winner!!

- Scalable, general purpose code for performing classical molecular dynamics (MD) simulations using highly accurate MGPT potentials

- MGPT semi-empirical potentials, based on a rigorous expansion of many body terms in the total energy, are needed in to quantitatively investigate dynamic behavior of d-shell and f-shell metals.

524 million atom simulations on 64K nodes achieved 101.5 TF/s sustained. **Superb strong and weak scaling** for full machine - (“very impressive machine” says PI)

Visualization of important scientific findings already achieved on BG/L: Molten Ta at 5000K demonstrates solidification during isothermal compression to 250 GPa
Rapid Resolidification in Tantalum: Excellent scaling of ddcMD on BG/L

- Nucleation is initiated at multiple independent sites in each sample cell
- Growth of solid grains initiates independently, but soon leads to grain boundaries which span the simulation cell
- 101.5 sustained teraflop/s achieved.

- The ddcMD team is currently using 131,072 CPUs of BG/L for unprecedented five hundred million atom MGPT simulations
Simulating the Kelvin-Helmholtz Instability

- Simulating KH instabilities via quasi 2D MD
  - micron length scales in two dimensions with a nanometer length scale in the third dimension
  - equal parts molten copper and aluminum separated by a planar interface
  - a shear was applied with a velocity difference of 2000 m/s

- 9 billion atoms simulated for over 1.4 nanoseconds
  - this represents a total of \textit{6.4 quadrillion} \((10^{15})\) position and velocity updates requiring \textit{36 quintillion} \((10^{18})\) flops
- Currently simulating 62.5 billion atoms in a fully 3D geometry – this represents a cubic micron of material
Modeling hydrodynamics at the atomic scale

- BlueGene/L upgrade enables unprecedented 12.0 µm x 6.0 µm quasi-2D simulation of 9 billion atoms developing a Kelvin-Helmholtz instability
- Currently performing tour-de-force simulation of 1 cubic micron of sheared molten metal (62.5 Billion atoms) - first ever in 3D
- Atomistic simulation paves the way for creation of fully-resolved sub-zone model
Qbox: First Principles Molecular Dynamics
Francois Gygi UCD, Erik Draeger, Martin Schulz, Bronis de Supinski, LLNL
Franz Franchetti Carnegie mellon, John Gunnels, Vernon Austel, Jim Sexton, IBM

- Treats electrons quantum mechanically
- Treats nuclei classically
- Developed at LLNL
- BG Supported provided by IBM
- Simulated 1,000 Mo atoms with 12,000 electrons
- Achieves 207.3 Teraflops sustained.
  - (56.8% of peak).

Qbox simulation of the transition from a molecular solid (top) to a quantum liquid (bottom) that is expected to occur in hydrogen under high pressure.
Qbox

<table>
<thead>
<tr>
<th></th>
<th># calls</th>
<th>avg. msg size (bytes)</th>
<th>time (s)</th>
<th>% MPI</th>
<th>% total</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPI_Bcast</td>
<td>4048</td>
<td>950681.5</td>
<td>401.0</td>
<td>70.2%</td>
<td>14.5%</td>
</tr>
<tr>
<td>MPI_Allreduce</td>
<td>54256</td>
<td>35446.1</td>
<td>103.6</td>
<td>18.1%</td>
<td>3.7%</td>
</tr>
<tr>
<td>MPI_Alltoallv</td>
<td>5635</td>
<td>399.1</td>
<td>41.8</td>
<td>7.3%</td>
<td>1.5%</td>
</tr>
<tr>
<td>MPI_Reduce</td>
<td>940</td>
<td>186007.1</td>
<td>15.2</td>
<td>2.7%</td>
<td>0.5%</td>
</tr>
<tr>
<td>MPI_BARRIER</td>
<td>2352</td>
<td>0.0</td>
<td>8.8</td>
<td>1.5%</td>
<td>0.3%</td>
</tr>
</tbody>
</table>
Large-scale Atomic/Molecular Massively Parallel Simulator


- LAMMPS has potentials for soft materials (biomolecules, polymers) and solid-state materials (metals, semiconductors) and coarse-grained or mesoscopic systems. It can be used to model atoms or, more generically, as a parallel particle simulator at the atomic, meso, or continuum scale.

- LAMMPS in HPC uses message-passing techniques and a spatial-decomposition of the simulation domain.

- LAMMPS is used as a Sequoia Benchmark. The EAM potential, weak scaling problem is used, as defined on the LAMMPS web site under “benchmarks”.

- **One Million MPI Task Application Demonstrated**
  - Demonstrated correct function with almost perfect scaling.
  - 64 MPI Tasks / Node, 256k Nodes.
Strategic Direction of CBC Research

at the intersection of Information Technology and Life Sciences

• Managing and Analyzing the Data Explosion
• Predicting Biological Outcomes with Scale of Computing
• Systems Biology to deal with Complexity

Predict Viral Mutations  
Ligand Binding  
Predict Toxicity  
Genomics
RHODOPSIN GPCR

Membrane Proteins – Cell Signalling, Ion/Nutrient Transport. Are targets of Many drugs.

Diseases associated with malfunction of GPCRs are:

Congestive Heart Failure, Hypertension, Stroke, Cancer, Ulcers, Allergies, Asthma, Anxiety, Psychosis, Migraines, Parkinson’s Disease.
Rhodopsin Core Hydration

Fig. 2. The counterion switch and internal hydration. (a) and (c) Time evolution of the distances between the ionone ring and Glu113 and Glu181 for the counterion switch and complex counterion models, respectively. (b) and (d) Number of water molecules inside the protein cavity for the counterion switch and complex counterion simulations, respectively. (d) The equivalent time series for a control simulation of dark-state rhodopsin. The number of internal water molecules at $t=0$ differs in the two simulations because the number of water molecules fluctuated during dark-state equilibration.

Fig. 3. Internal water molecules solvate retinal during MI formation. A snapshot from the last point of the complex counterion simulation is shown, depicting the large number of water molecules inside the protein cavity and their interaction with the retinal. Trp265 (green) has moved away from the retinal (yellow) head group. The image was created using PyMOL.26

2012/4/25
Computation Rates

![Graph showing computation rates for various processes and models. The x-axis represents Atoms/Node, ranging from 0 to 0.1. The y-axis represents Computation Rate (time-steps/second), ranging from 0 to 1400. The legend includes labels for different processes and models such as Hairpin SPI 64^3 (V5), SOPE SPI 64^$ (V5), Hairpin SPI 64^3 (V4), SOPE SPI (V5), Rhodopsin SPI (V5), SOPE SPI (V4), ApoA1 SPI (V5), Rhodopsin SPI (V4), ApoA1 SPI (V4), SOPE MPI (V4), Rhodopsin MPI (V4), ApoA1 MPI (V4), ApoA1 NAMD Msging Layer, ApoA1 NAMD MPI. The graph shows varying computation rates across different processes and models, with some processes achieving higher rates than others.]
Rhodopsin: Nanoseconds to Microseconds

- Prior to Blue Matter on BG/L, longest simulation of Rhodopsin: 40ns
- 512 node BG/L prototype & Blue Matter: 1 x 120ns (JACS 2005)
- 20 rack BG/L Watson & Blue Matter: 26 x 120 ns (PNAS 2006)
- 20 rack BG/L Watson & Blue Matter: 1 x 2500 ns + several 1000+ ns
- One month for microsecond simulations of biologically interesting systems
Influenza: Research Challenges

Computational modeling of mutations in proteins of Influenza A(H5N1) and A(H3N2)

Goal: Predict biological consequence of mutations

**Receptor binding** - specificity to avian v. human host

**Antibody binding** - escape from antibody neutralization

---

**Predicted Mutations: Receptor Binding**

<table>
<thead>
<tr>
<th>H5 HA mutation</th>
<th>α-2,3- (avian)</th>
<th>α-2,6- (human)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ΔG_h</td>
<td>ΔG_f</td>
</tr>
<tr>
<td>V13SS</td>
<td>5.68 (0.58)</td>
<td>4.56 (0.4)</td>
</tr>
<tr>
<td>A138S</td>
<td>4.39 (0.1)</td>
<td>4.52 (0.58)</td>
</tr>
<tr>
<td>V13SS + A138S</td>
<td>9.48 (0.8)</td>
<td>8.64 (0.63)</td>
</tr>
</tbody>
</table>

**Prediction of Ab Escape Mutants**

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Calc. ΔΔG (kcal/mol)</th>
<th>Expt. ΔΔG (kcal/mol)</th>
<th>Mutation</th>
<th>Calc. ΔΔG (kcal/mol)</th>
<th>Expt. ΔΔG (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T131I</td>
<td>5.20 ± 0.94</td>
<td>5.0</td>
<td>T131F</td>
<td>5.68 ± 1.48</td>
<td></td>
</tr>
<tr>
<td>T131G</td>
<td>-3.72 ± 0.69</td>
<td></td>
<td>T131W</td>
<td>7.46 ± 1.51</td>
<td></td>
</tr>
<tr>
<td>T131A</td>
<td>-2.81 ± 0.91</td>
<td></td>
<td>T131L</td>
<td>3.15 ± 1.19</td>
<td></td>
</tr>
<tr>
<td>T131C</td>
<td>0.117 ± 1.24</td>
<td></td>
<td>T131H</td>
<td>3.84 ± 1.17</td>
<td></td>
</tr>
<tr>
<td>T131V</td>
<td>2.58 ± 0.89</td>
<td></td>
<td>T131Y</td>
<td>6.01 ± 1.31</td>
<td></td>
</tr>
<tr>
<td>T131M</td>
<td>0.57 ± 1.63</td>
<td></td>
<td>T131N</td>
<td>2.92 ± 1.16</td>
<td></td>
</tr>
<tr>
<td>T131Q</td>
<td>1.22 ± 1.20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T131S</td>
<td>-0.48 ± 1.57</td>
<td></td>
<td>S157L</td>
<td>4.10 ± 1.69</td>
<td>3.7</td>
</tr>
</tbody>
</table>
IBM Computational Biology Center

Computational Medicine

Predictive Toxicology

Multiscale model of organs

Computational modeling of cardiac disease - potential for personalized medicine.

Fred Hoyle postulated triple-alpha process for nucleosynthesis beyond Be inside stars.

Experimentally verified in 1954, but attempts to calculate the Hoyle state have been unsuccessful since 1954.

Hoyle state was simulated successfully (May 9, 2011) with 1 week calculation time on JUGENE.
QCD on the Lattice
Strong Scaling

- Test on 72 racks BG/P installation at supercomputer center Jülich (Gerhold, Herdioza, Urbach, K.J.)
- using tmHMC code (Urbach, K.J.)
KEK: Nuclear Force from Lattice QCD
Seismology, Oil Exploration
Earthquake Simulation Example: AWM PetaShake Benchmark Run on Blue Gene

- 800km x 400km x 100km (V4)
- Spatial resolution = 100 m
- 32 billion mesh points
- Min. S wave velocity = 500 m/s
- Simulated time = 1 sec
- Time steps = 101
- Single point source

A breakthrough in the field of earthquake ground motion simulation

Seismic hazard map of Southern California, showing four proposed regions for PetaSHA simulations: Northridge domain (V1), PSHA site volume (V2), regional M 7.7 domain (V3), and regional M 8.1 domain (V4)
Focus on Upstream Processing

- Seismic Processing
  - Converting Raw Field Data to 3D earth subsurface images
  - Extremely parallel (HPC, GPU…)

- Reservoir Simulation
  - Economic/technical model of reservoir to determine drill decision and planning
Petroleum Industry Applications on the Blue Gene

- **ExxonMobil Upstream Research**
  - Electrical Method for deep sub-salt imaging.
  - Node performance near parity with 2.8 GHz Xeon processors
  - 16 rack run for 27 hours (32,768 cores)
  - Run is roughly equivalent to a 28,000 cores Xeon cluster (which doesn’t exist, and if it did would consume roughly 5x floor space and power)
  - Equivalent run on a standard Intel 1024 node (dual-processor nodes) Linux Cluster would take 16 days (assuming all nodes could stay operational for that long)
  - See Publication – Commer, et.al. (2008)*

- **Tsunami Development (ISV)**
  - Kirchoff-based PSTM and PSDM seismic imaging apps
  - Port complete, production tests to 4096 processors
  - High Resolution Seismic Imaging
  - Process 300 km²/day/rack PSTM, 41Million traces.
  - Tsunami states that one Blue Gene /L rack equals approximately a 1500 processor Opteron cluster

Application Performance, LLNL and ANL

- Applications running up to 16 racks, 16k processors, 256k cores
  - more performance to come.
- Linpack – SOW Target Achieved
  - 84.3% on 1 Rack, 78.8% on 16 Racks.
  - 85%+ of peak may be achievable with careful power management.
- LLNL Sequoia Apps – SOW Targets Achieved.
  - Base Apps: 154% (was 148%) of committed target
  - Tuned Apps: 121% (was 124%) of committed target
- ANL Mira Apps – SOW Targets On Track
  - Significant progress in last 3 months
  - 4 Apps at 10x speed up per node, P to Q, meets SOW
  - 2 Apps at 8x speed-up per node, P to Q, need 3.
  - Expect to achieve SOW targets in next quarter.
- One Million MPI Task Application Demonstrated
  - Demonstrated correct function with almost perfect scaling of LAMMPS (Molecular Dynamics).
  - 64 MPI Tasks / Node, 256k Nodes.
Summary

- Emerging HPC landscape is extremely complex
- A time of extraordinary potential
  - Game-changing capability is now available
- A time of significant challenge
  - Just as HPC starts to have real scientific and industrial impact - it gets extraordinarily hard.
- A radical research and development approach required
  - Multidisciplinary from domain to system design
  - Collaboration across research teams within universities
  - Collaboration between universities and industry