SAND2015-4229C

Exceptional service in the national interest

Sandia National Laboratories

Enabling Performance Portability for Large-scale Atomistic Simulations on Next-Generation Computing Architectures

Ray Shan

Sandia National Laboratories

Albuquerque, New Mexico

CAMS Workshop

Gainesville, FL

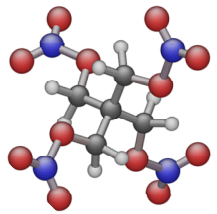
U.S. DEPARTMENT OF ENERGY

National Nuclear Security Administration

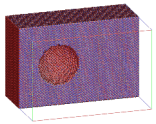
Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND NO. 2015-

Large-scale simulations of materials

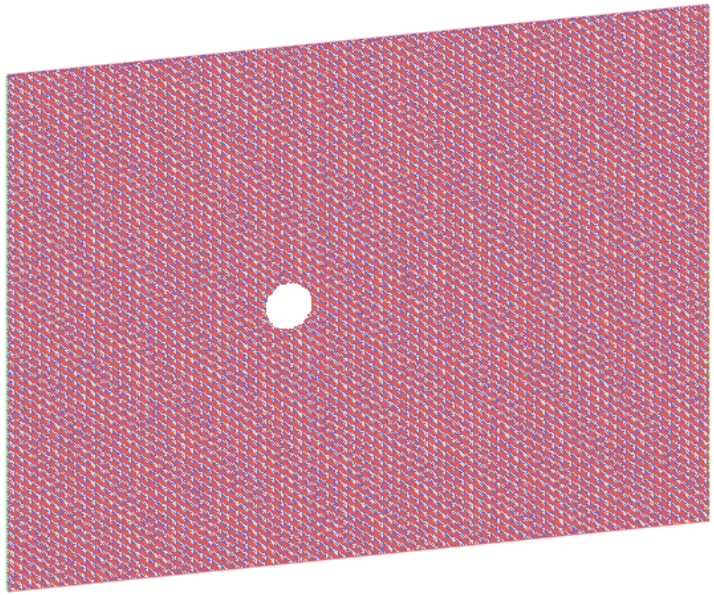
- Energetic materials (explosives)
  - Nano-structured defect enhanced chemical reaction initiation and detonation



PETN;
containing
C, H, O and N



$60 \times 40 \times 40$ nm^3
20 nm spherical void
8.9 million atoms
154,000 molecules



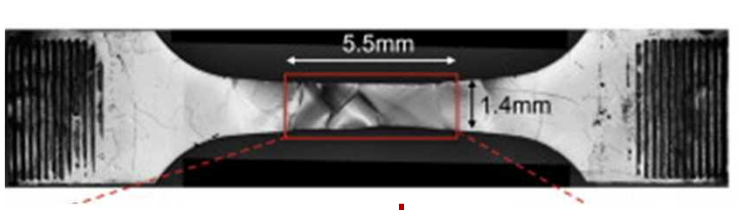
$300 \times 200 \times 1.3$ nm^3
20 nm cylindrical void
8.5 million atoms

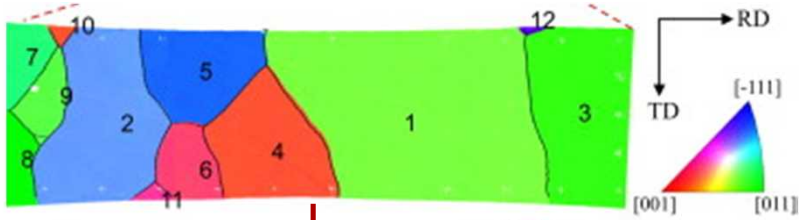
- T.-R. Shan, et. al., Proc. 15th International Detonation Symposium, accepted for publication (2015)
- T.-R. Shan, et. al., J. Phys.: Conf. Ser. 500, 172009 (2014)
- SAND 2014-1297C; SAND2015-1243C

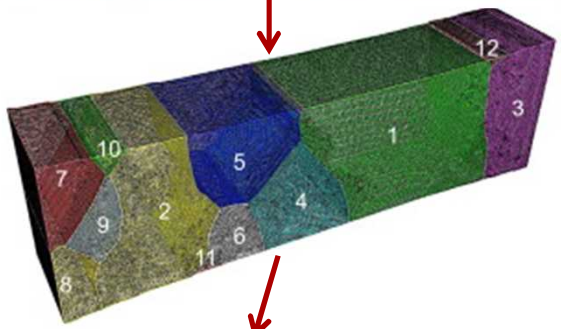
0.1 μm

Large-scale simulations of materials

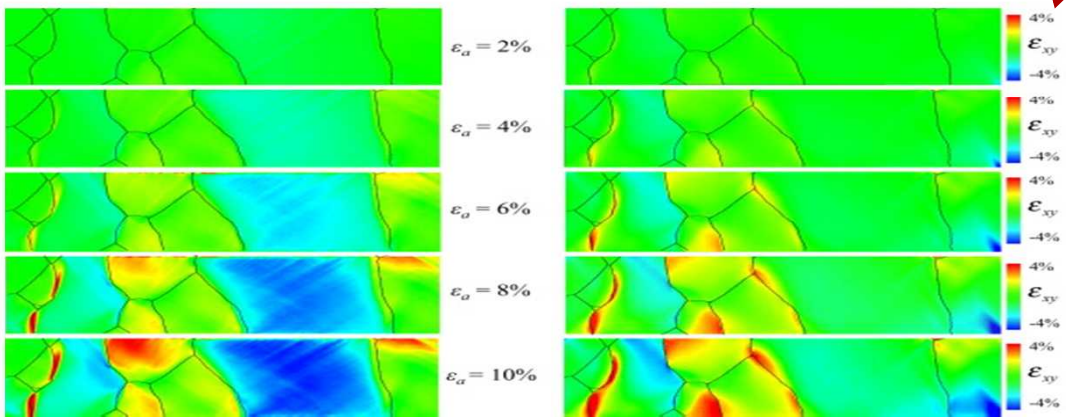
- Poly- and oligo- crystal metals
  - Plastic deformation and fracture mechanism







Hojun Lim, et. al., Int. J. Mech. Sci. 92 (2015)



- Atomistic simulations provide:
  - Yield stress;
  - Dislocation motion activation energy;
  - Grain boundary configurations

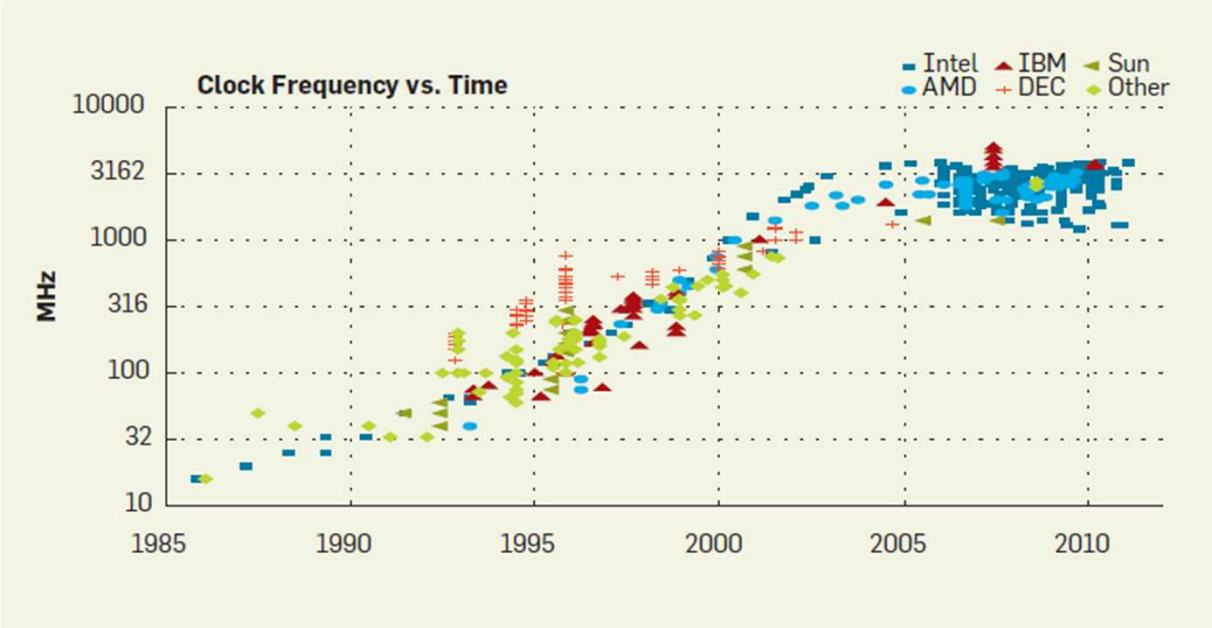
Large-scale simulations of materials

Sandia National Laboratories

- The need for large-scale atomistic simulations
  - In many other fields
    - Surface chemistry
    - Soft materials
    - Geomaterials
  - What is the role of theoretical and computational methods in the design, processing, and application of materials?
  - How can we better integrate the latest computational approaches with experimental data to improve predictability?
- We need to exploit the power of large-scale supercomputers
  - Increased thread counts
  - Decreased memory per thread
  - Architecture specific memory access patterns

Next-generation supercomputers

- Central processing unit (CPU) clock speed
  - Moore's law … till 2005



cacm.acm.org

  - Because of "power wall"!

  - If we can't make CPUs run faster in a supercomputer …

Next-generation supercomputers

Sandia National Laboratories

- We make many more CPUs that run slower!
- Next-generation computing architectures
  - Multi-threading
    - IBM PowerPC CPU
      - BlueGene/Q: 1.6 GHz cores, 16 cores/node, 4 threads/core
      - **64** parallel processes/node
  - Co-processors (accelerators)
    - Intel Many Integrated Core (MIC) architecture
      - Xeon Phi series (Knights Corner, Knights Landing,, Knights Hill)
      - 57 – 61 cores (1.0 – 1.2 GHz each), 4 threads/core = **244** processes/accelerator
    - NVIDIA general purpose graphics processing unit (GPGPU)
      - Kepler series (K10, K20, K20x, K40, K80)
      - 1 – 2 cores (562 – 745 MHz each), 1536 – 2880 threads/core
      - **1536 – 4992** parallel processes/accelerator

Next-generation supercomputers

Sandia National Laboratories

- Intel Xeon Phi co-processors

#1: Tianhe-2 (China)
3,120K cores, 33.8 PFLOPS

#7: Stampede (US)
462K cores, 5.1 PFLOPS

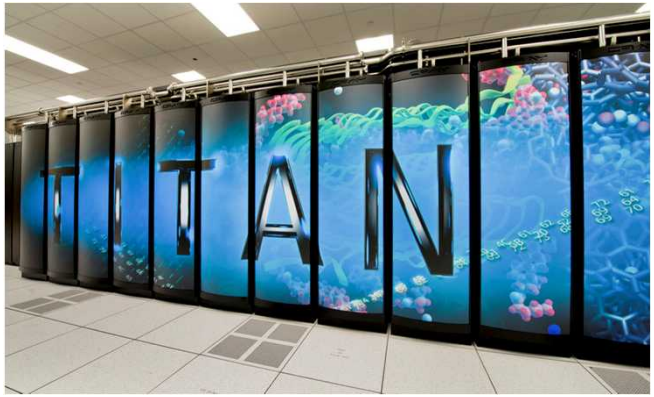
- PFLOPS = peta FLOPS = 10^{15} FLOPS = 1 thousand trillion FLOPS
- 2.5 GHz CPU = 10^{10} FLOPS

Next-generation supercomputers

Sandia National Laboratories

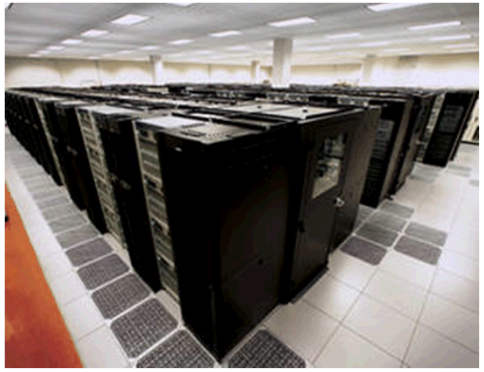
- Nvidia GPGPU co-processors

#2: Titan, (US)
560K cores, 17.5 PFLOPS



#6: Piz Daint (Switzerland)
115K cores, 6.2 PFLOPS

#10: Undisclosed site (US)
722K cores, 3.5 PFLOPS



Next-generation supercomputers

Sandia National Laboratories

- IBM BlueBene/Q PowerPC SMT processors

#3: Sequoia, (US)
1,572K cores, 17.1 PFLOPS

#5: Mira, (US)
782K cores, 8.5 PFLOPS

#6: JUQUEEN (Germany)
458K cores, 5.0 PFLOPS

#9: Vulcan (US)
393K cores, 4.2 PFLOPS

Next-generation supercomputers

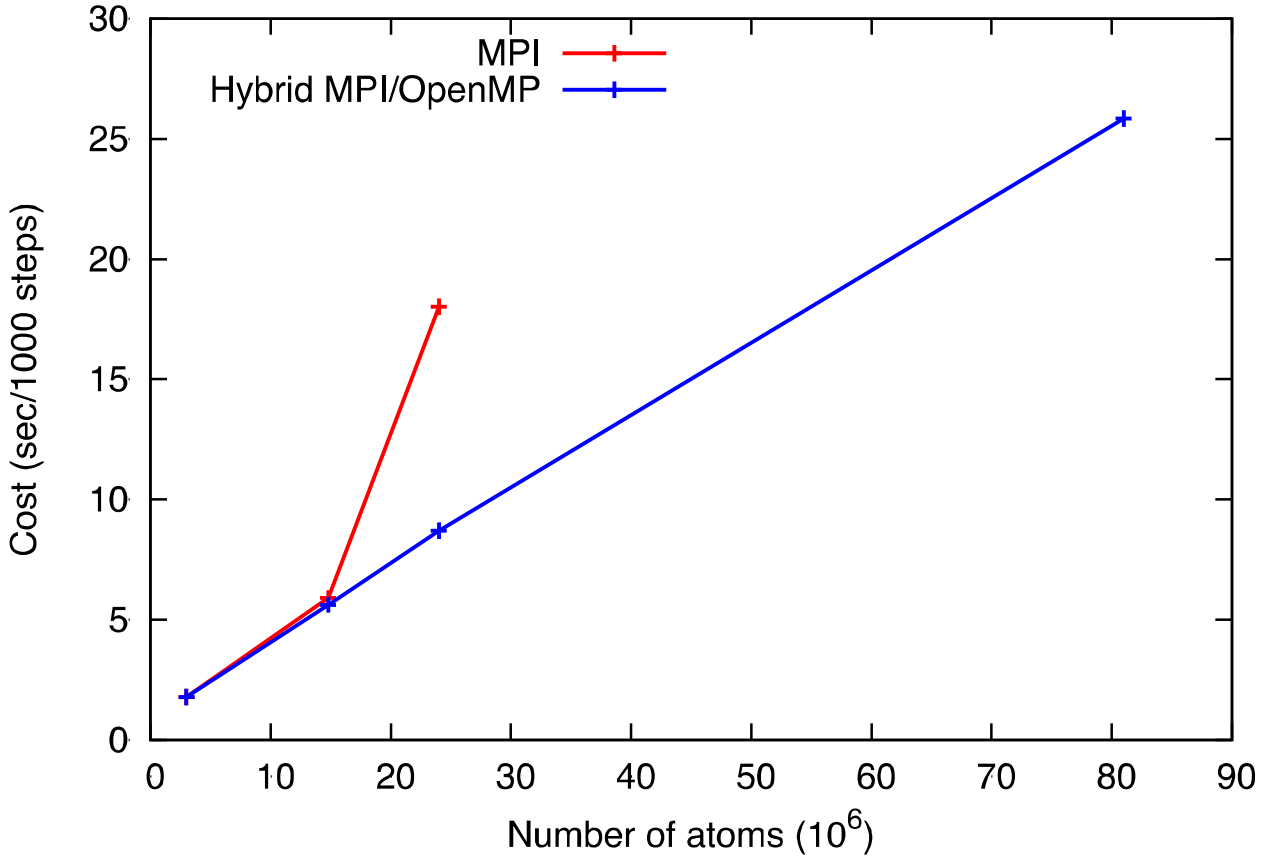
Sandia National Laboratories

- Many-core architectures
  - 9/10 of the top 10 supercomputers are many-core architectures
  - With more to come, 150-300 PFLOPS/s
    - Trinity at SNL and LANL (Intel Xeon Phi 3rd gen. KNH), coming 2016
    - Sierra at LLNL (IBM PowerPC SMT), coming 2017
    - Summit at ORNL (Nvidia GPU), coming 2018
  - Domination of many-core architectures

- What kind of parallelism works on many-core?
  - Distributed memory (MPI)
  - Shared memory
  - Hybrid distributed/shared memory

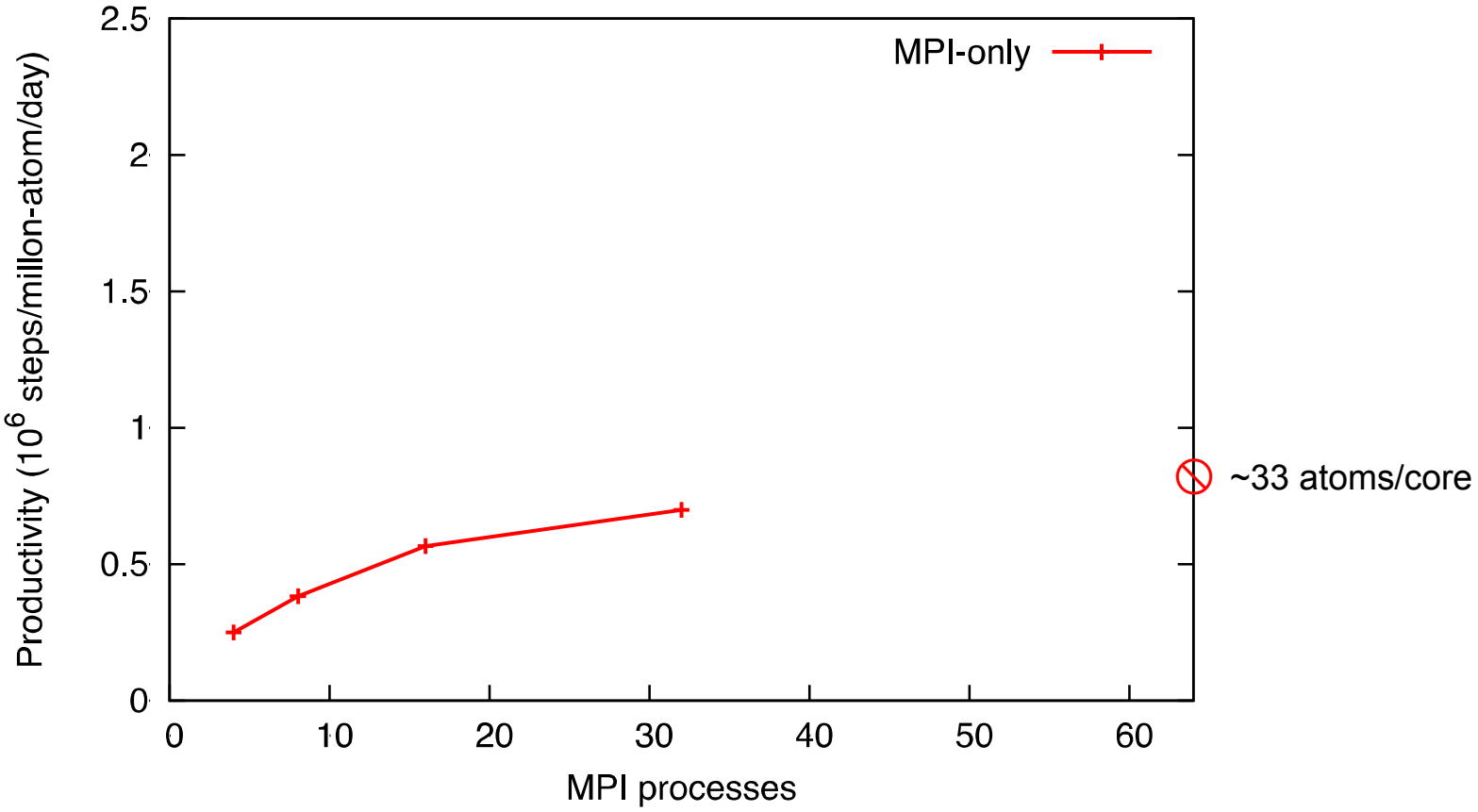
MPI-only parallelism

- CHARMM scaling of a water benchmark
  - 512 IBM BG/Q (Mira) nodes, up to 32,768 tasks



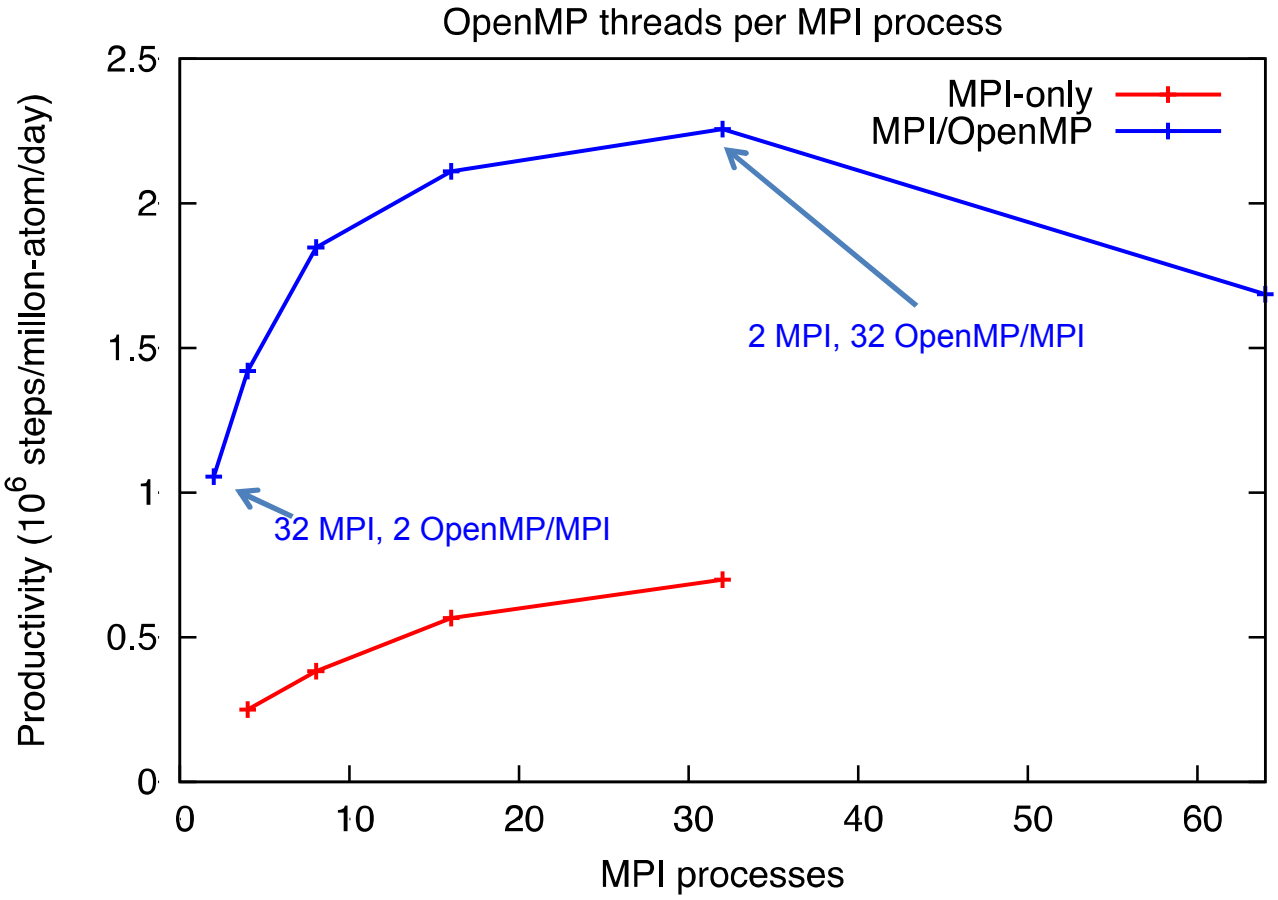
MPI-only parallelism

- ReaxFF scaling of a 8.6 million atom PETN structure
  - 4,096 IBM BG/Q (Mira) nodes, up to 262,144 tasks



MPI-only parallelism

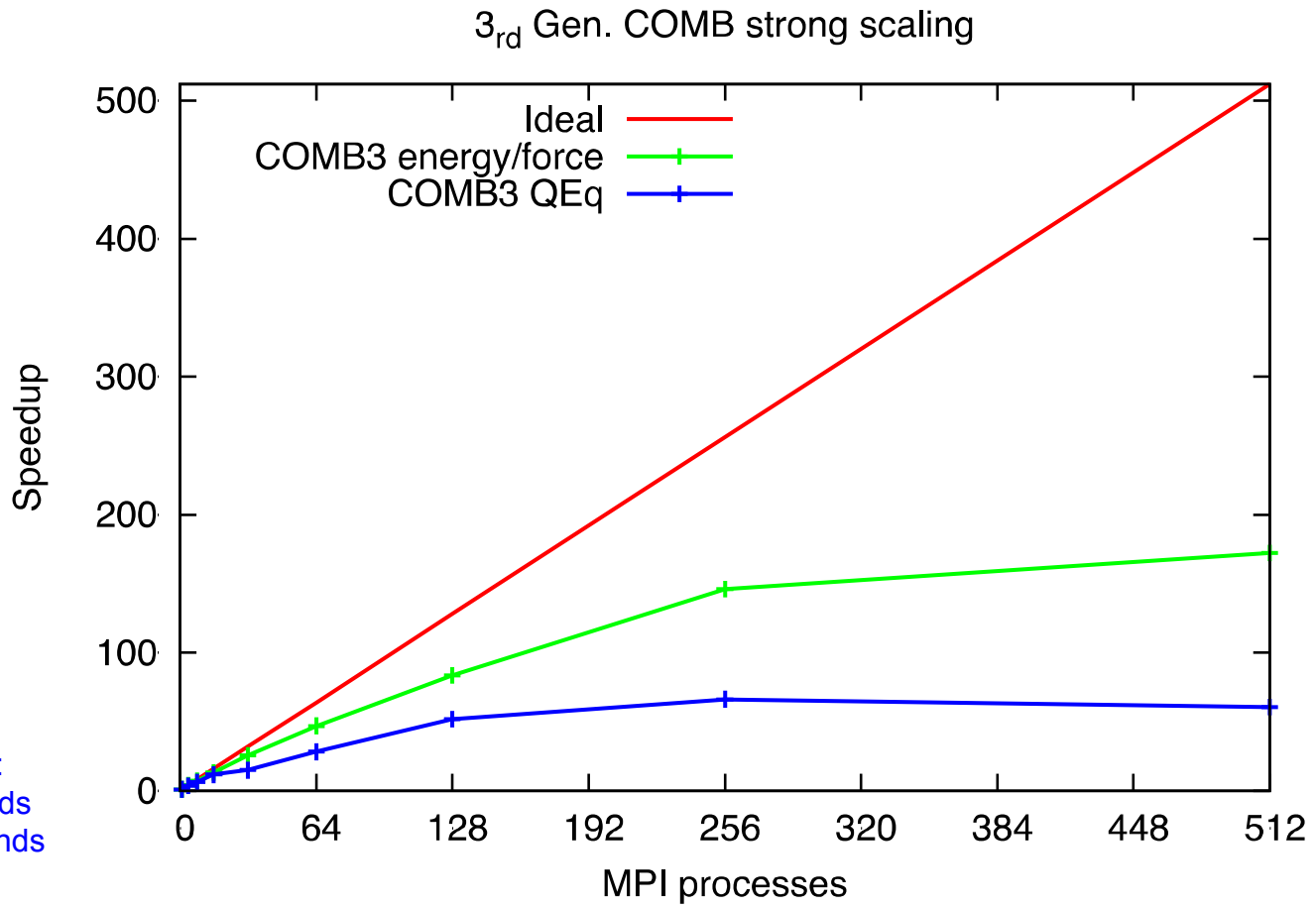
- ReaxFF scaling of a 8.6 million atom PETN structure
  - 4,096 IBM BG/Q (Mira) nodes, up to 262,144 tasks



Aktulga, Coffman, Jiang, Knight, Shan, *in preparation* (2015)

MPI-only parallelism

- COMB3 strong scaling for a 10,912 atom Cu_2O-CHO system



Timing on 1 MPI:
E/F: 184 seconds
QEq: 718 seconds

Timing on 512 MPI:
E/F: 1.06 seconds
QEq: 11.8 seconds

Beyond MPI-only parallelism

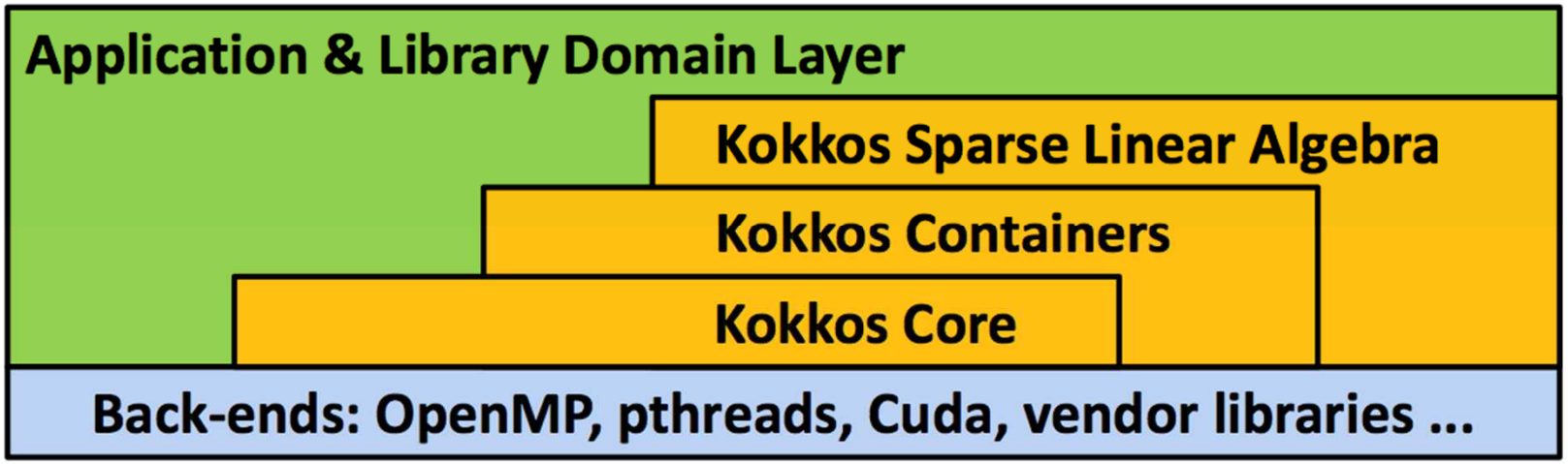
Sandia National Laboratories

- We obviously need parallelism beyond MPI-only
- Hybrid MPI/OpenMP parallelism works well on multi-threaded architecture, i.e., IBM BG/Q
  - From simple pairwise to the most complicated variable charge, reactive potentials
  - But not on other many-core architecture, i.e., Intel MIC and NVIDIA GPGPU
- We can't afford the time and labor to port our science codes to the various computing architectures
- Is there a universal, **one size fits all** library/programming interface that works for all architectures?

Kokkos: Introduction

Sandia National Laboratories

- Yes! Kokkos *is* the one size fits all solution!
- What is Kokkos?
  - Sandia developed library/programming interface
  - Enables performance portability across many-core devices
  - A collection of C++ libraries
    - Standard C++
    - Not a new language nor a language extension (OpenMP, OpenCL, CUDA, etc)

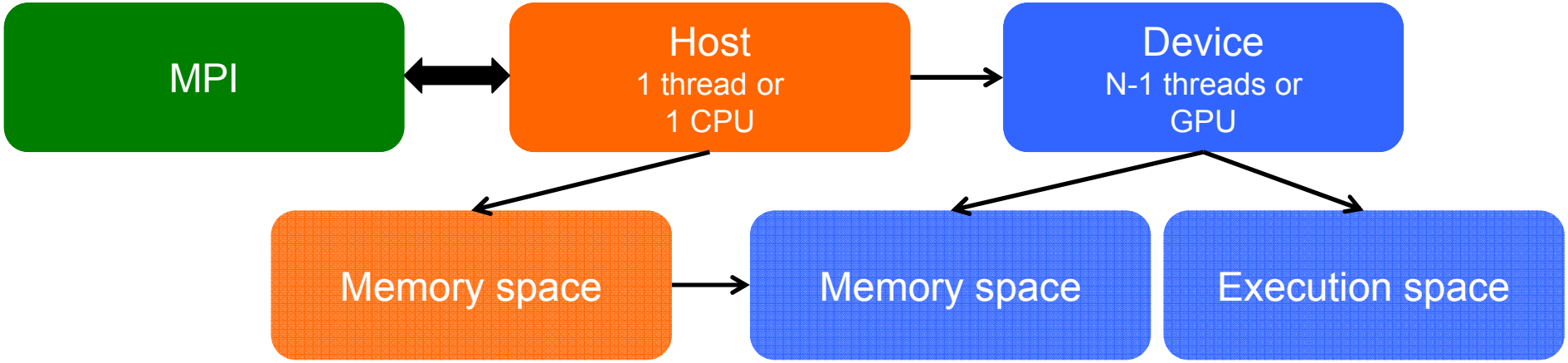


SAND2014-2317C

Kokkos: Introduction

Sandia National Laboratories

- Two fundamental abstractions
  - Host and Device spaces



  - Multi-dimensional array management
    - No prescribed *layout* (map between index and the set of values)
    - Map multi-index (I,j,k,…) to device's memory space
      - Efficient: index computation and memory use
    - Device-specific memory access pattern chosen at compile time

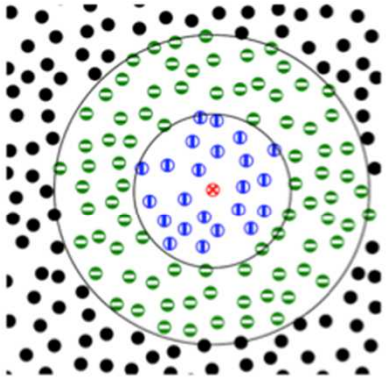
SAND2014-2317C

Kokkos: Sample code

- Lennard-Jones force calculation
  - Solve Newton's equations of motion for N particles
  - Simple Lennard-Jones force expression:
  - Use neighbor list to avoid N^2 loops

$$F_i = \sum_{j, r_{ij} < r_{cut}} 6\varepsilon \left[\left(\frac{\varsigma}{r_{ij}} \right)^7 - 2 \left(\frac{\varsigma}{r_{ij}} \right)^{13} \right]$$

```
pos_i = pos(i);
for( jj = 0; jj < num_neighbors(i); jj++) {
  j = neighbors(i,jj);
  r_ij = pos_i - pos(j); //random read 3 floats
  if ( |r_ij| < r_cut )
    f_i += 6*e*( (s/r_ij)^7 - 2*(s/r_ij)^13 )
}
f(i) = f_i;
```

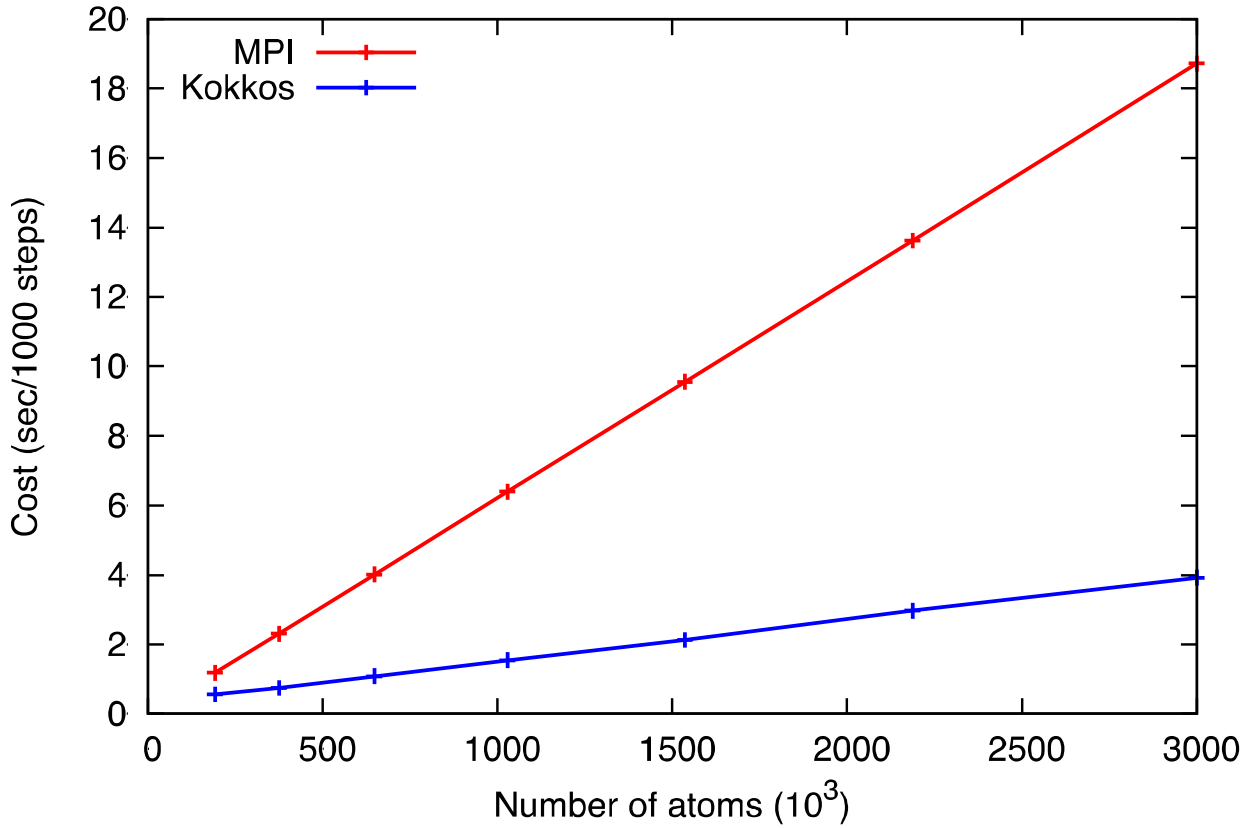


- Standard C/C++ language

SAND2014-2317C

Kokkos: Scaling

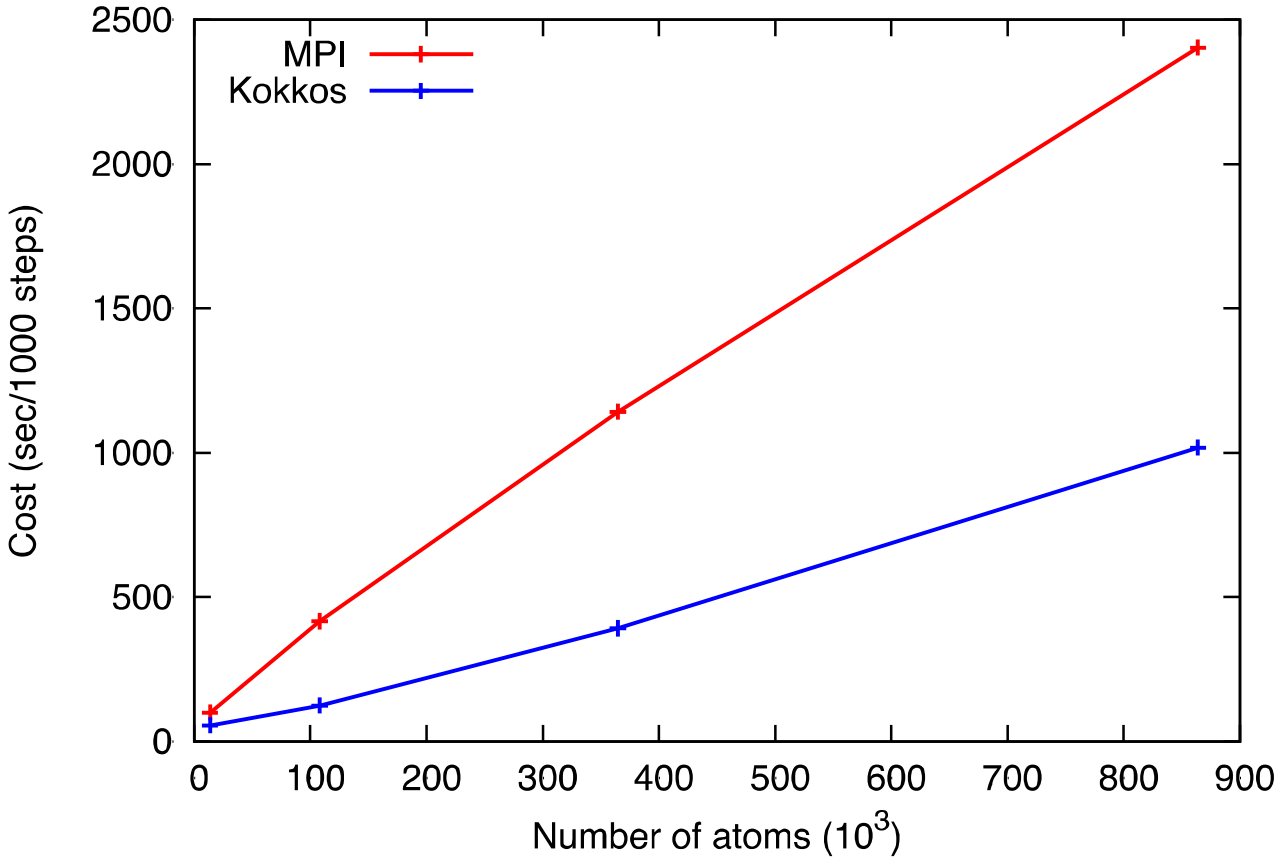
- CHARMM scaling of a water
  - 1 Shannon node (2 8-core Intel Xeon + **2 NVIDIA Tesla K20x**)



Kokkos: Scaling

Sandia National Laboratories

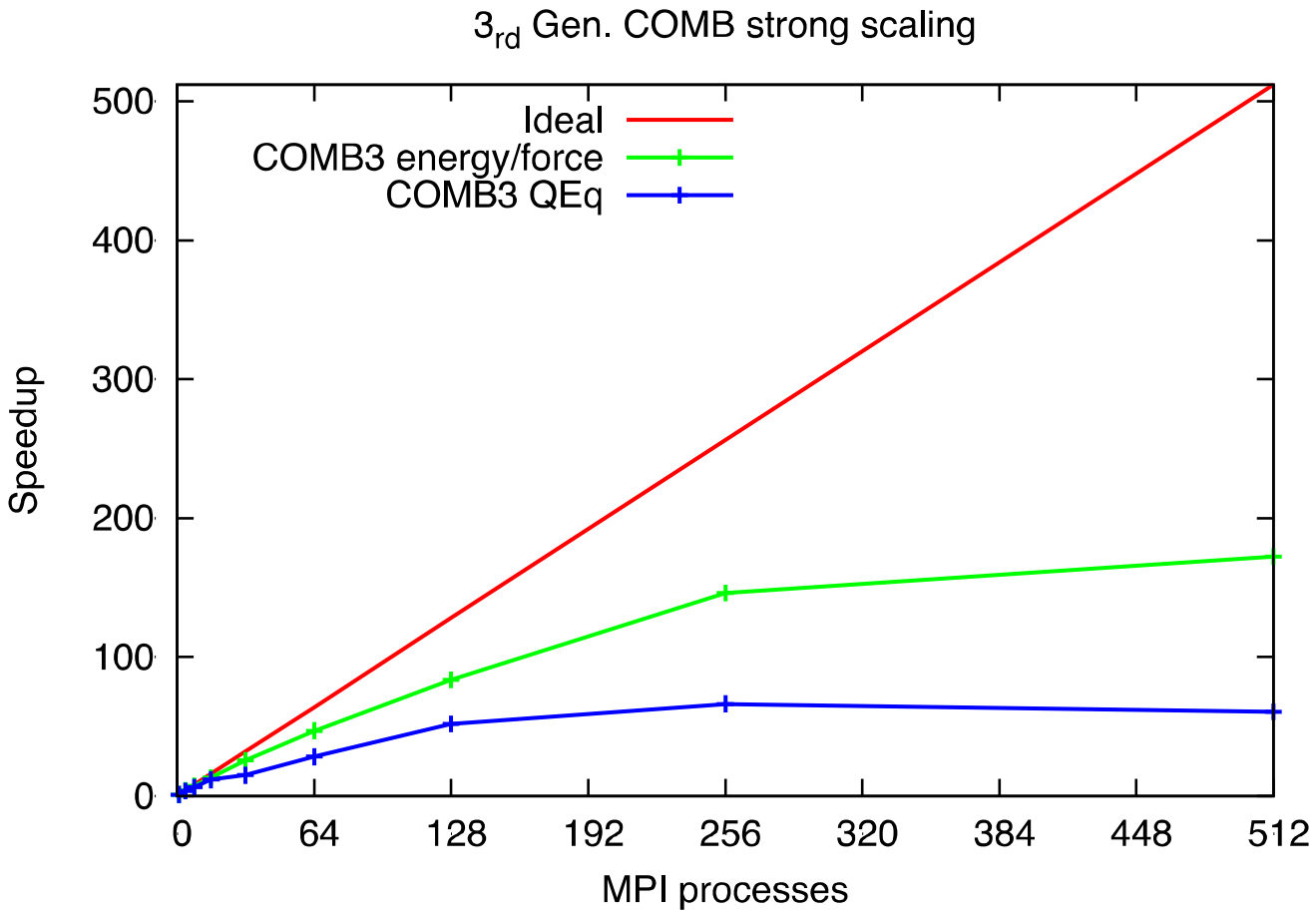
- ReaxFF scaling of a SiO_2 benchmark *
  - 4 Shannon nodes (2 8-core Intel Xeon + **2 NVIDIA Tesla K20x**)



* Includes only bond order, bond energy, Coulomb, and LJ interactions

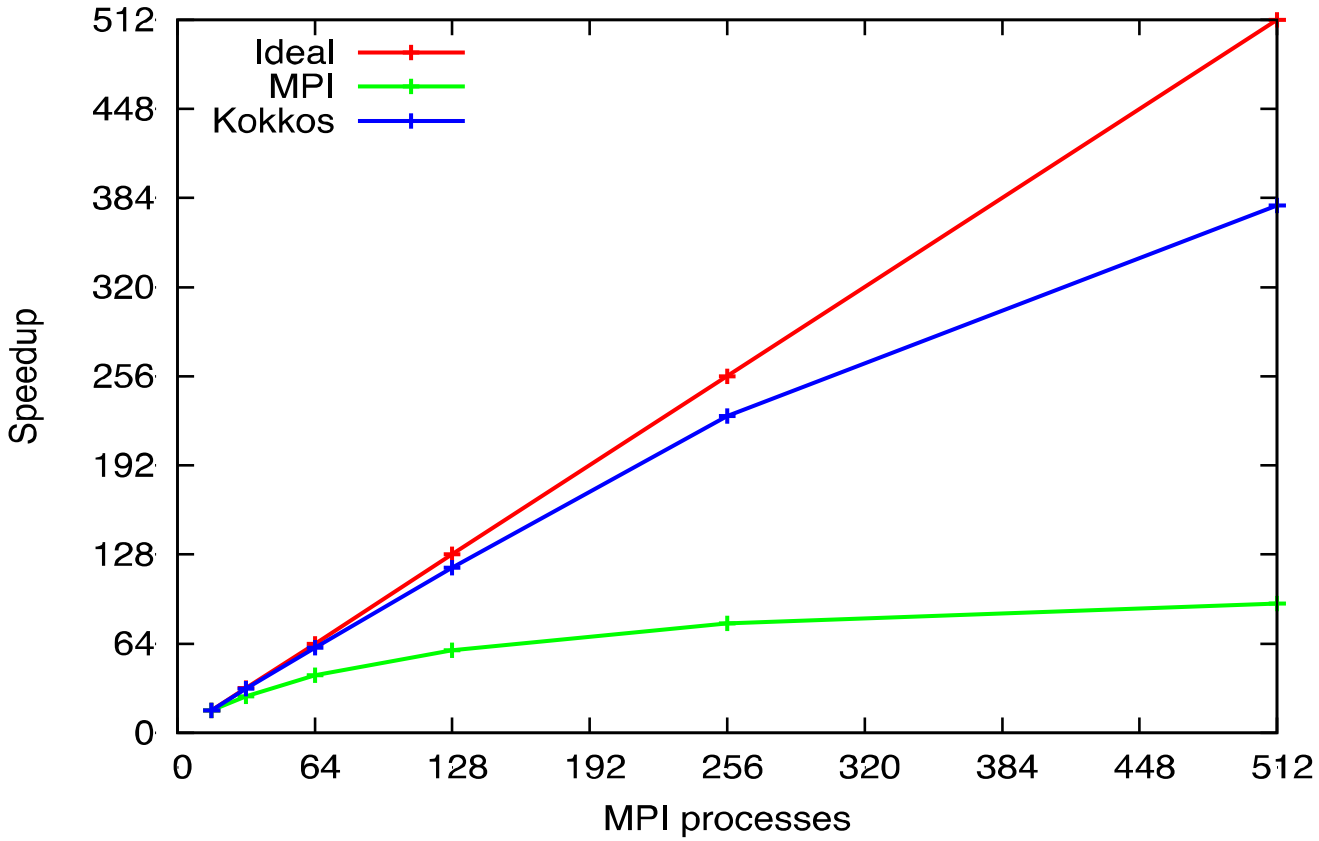
Kokkos: Scaling

- COMB3 strong scaling for a 10,912 atom Cu_2O-CHO system



Kokkos: Scaling

- ReaxFF strong scaling for a 10,940 atom SiO_2 benchmark *
  - 32 Compton nodes (2 8-core Intel Xeon + **2 Intel Xeon Phi**)



* Includes only bond order, bond energy, Coulomb, and LJ interactions (~60%)

Kokkos: current and future plans

Sandia National Laboratories

- LAMMPS
  - Most LJ variants, EAM, SW, and Tersoff potentials
  - Various bond, angle, dihedral styles
- Trillinos package
- MiniMD and MiniFE
- Other Sandia software
  - CTH, Alegra
- Other external software

Summary

Sandia National Laboratories

- Large-scale atomistic simulations needed in many fields to solve interesting and important problems
- Current dominant supercomputers are many-core architectures
- Kokkos enables performance portability across architectures
  - *All the rides are included in the admission price!*

Concluding remarks

- Current top 10 supercomputers are petascale (10^{15}) machines
- Exascale supercomputer expected by 2020-2023
  - > 1 EFLOPS; 10^{18} FLOPS; 1 million trillion FLOPS

RANK	SITE	SYSTEM	CORES	RMAX (TFLOP/S)	RPEAK (TFLOP/S)	POWER (KW)
1	National Super Computer Center in Guangzhou China	**Tianhe-2 (MilkyWay-2)** - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P NUDT	3,120,000	33,862.7	54,902.4	17,808
2	DOE/SC/Oak Ridge National Laboratory United States	**Titan** - Cray XK7 , Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x Cray Inc.	560,640	17,590.0	27,112.5	8,209
3	DOE/NNSA/LLNL United States	**Sequoia** - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom IBM	1,572,864	17,173.2	20,132.7	7,890
4	RIKEN Advanced Institute for Computational Science (AICS) Japan	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect Fujitsu	705,024	10,510.0	11,280.4	12,660
5	DOE/SC/Argonne National Laboratory United States	**Mira** - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM	786,432	8,586.6	10,066.3	3,945
6	Swiss National Supercomputing Centre (CSCS) Switzerland	**Piz Daint** - Cray XC30, Xeon E5-2670 8C 2.600GHz, Aries interconnect , NVIDIA K20x Cray Inc.	115,984	6,271.0	7,788.9	2,325
7	Texas Advanced Computing Center/Univ. of Texas United States	**Stampede** - PowerEdge C8220, Xeon E5-2680 8C 2.700GHz, Infiniband FDR, Intel Xeon Phi SE10P Dell	462,462	5,168.1	8,520.1	4,510
8	Forschungszentrum Juelich (FZJ) Germany	**JUQUEEN** - BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect IBM	458,752	5,008.9	5,872.0	2,301
9	DOE/NNSA/LLNL United States	**Vulcan** - BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect IBM	393,216	4,293.3	5,033.2	1,972
10	Government United States	Cray CS-Storm, Intel Xeon E5-2660v2 10C 2.2GHz, Infiniband FDR, Nvidia K40 Cray Inc.	72,800	3,577.0	6,131.8	1,499

Concluding remarks

- In June of 2000 ASCI Red (Sandia) made history
  - First machine ever to achieve 1 TFLOP (10^{12})
  - $46,000,000
  - 1,600 sqft
  - 850 kW

Concluding remarks

- In June of 2000 ASCI Red (Sandia) made history
  - First machine ever to achieve 1 TFLOP (10^{12})
  - $46,000,000
  - 1,600 sqft
  - 850 kW

- In November, 2013 PlayStation 4 released
  - $400
  - 183 grams
  - 250 W
  - Computing power?

1.83 TFLOPS!

Concluding remarks

Sandia National Laboratories

- IBM Roadrunner
  - First PFLOP machine (1.04 PFLOP)
  - Operational in 2008
  - $100,000,000
  - 6000 sqft

- What is going to happen 13 years later, or roughly 2020?
  - Personal PFLOP workstaions!
  - < $1,000
  - Sits on your desk
  - Offers the same computing power as the $100M IBM Roadrunner

Concluding remarks

Sandia National Laboratories

"… Once we have an exascale machine, how small will a petaflop machine be? Will it fit in a backpack or under my desk? What is the research that a student could do that they can't do today?"

— Sumit Gupta
General manager for the Tesla accelerated computing business at Nvidia

- Exascale supercomputers or petascale personal workstations
  - Many-core architectures
  - Need to prepare ourselves for this near future

Acknowledgements

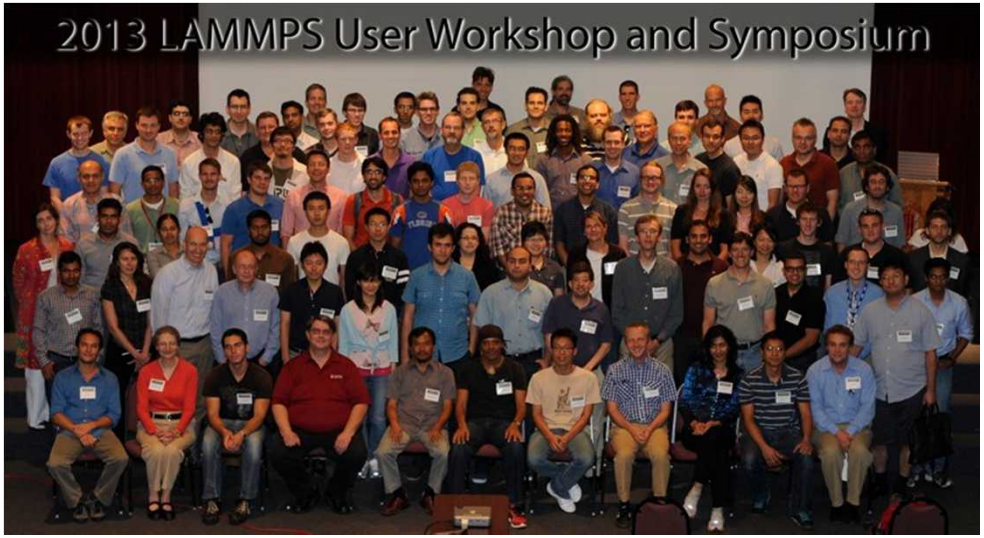
- LAMMPS development team
  - Steve Plimpton
  - Aidan Thompson
  - Stan Moore
- Kokkos development team
  - Carter Edwards
  - Christian Trott

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000

LAMMPS Users' Workshop

Sandia National Laboratories

- 2015 LAMMPS Users' Workshop and Symposium
  - August 5-7, 2015, Albuquerque, NM
  - Includes beginner's tutorials session, technical sessions, break-out sessions, and poster sessions
  - Includes invited and contributed talks/posters (deadline 6/22)
  - No registration fee; no registration deadline
  - http://lammps.sandia.gov/workshops.html



Questions

Sandia National Laboratories

- What is the role of theoretical and computational methods in the design, processing, and application of materials?
- How can we better integrate the latest computational approaches with experimental data to improve predictability?
- To what extent are computational methodologies available that are applicable to the *physics of interest* in actual systems (materials, length and time scales)? Where are the areas of greatest need for next-generation methodologies?
- How do we ensure the next generation of scientists and engineers can work with the latest computational methodologies and avoid insidious errors?
- Is there a universal approach for determining accurate error bars for the results of theoretical/computational methods?