Exceptional service in the national interest

The Mantevo Project : Tools for codesign

PSAAP II kickoff meeting

December 9-10, 2013

Albuquerque, NM

Richard Barrett
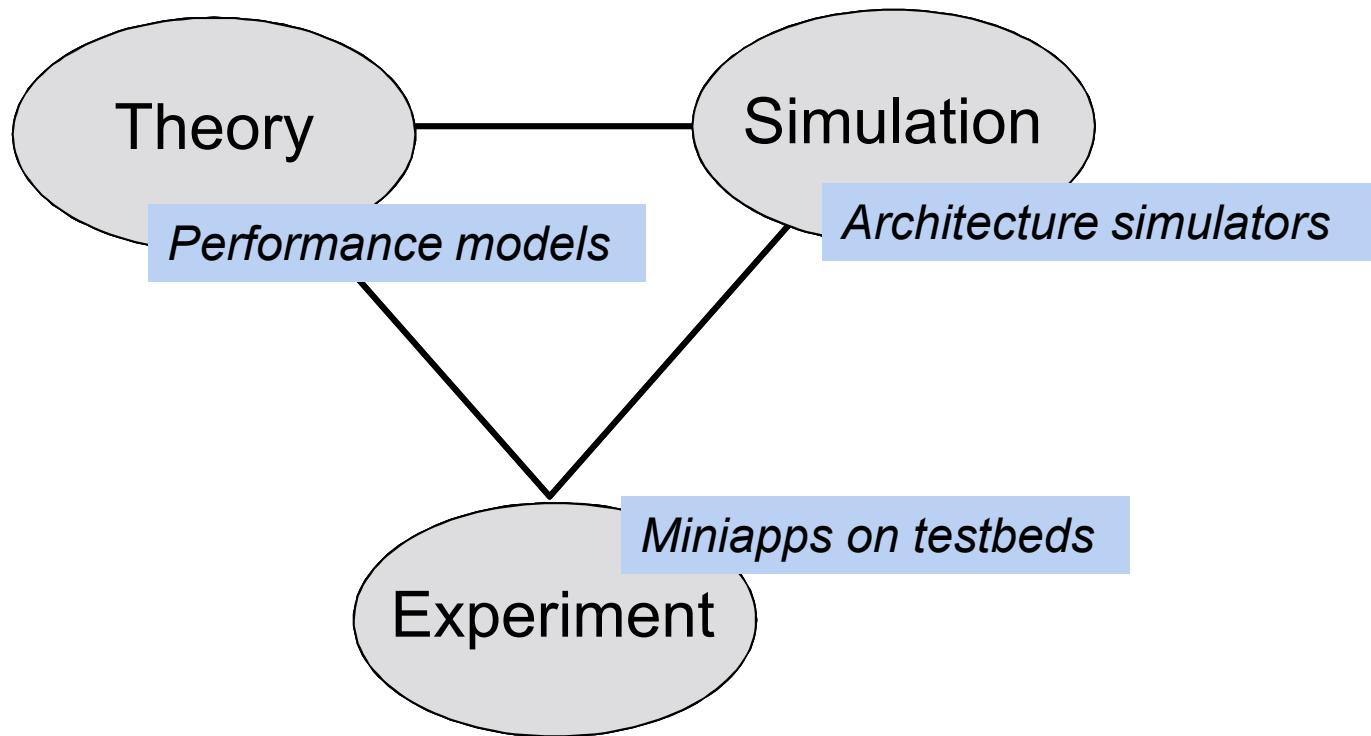Center for Computing Research
Sandia National Laboratories

Development Team

Daniel Barnette[1], Richard Barrett[1], David Beckingsale[2], Jim Belak[3], Mike Boulton[4], Paul Crozier[1], Doug Doerfler[1], Carter Edwards[1], Wayne Gaudin[5], Tim Germann[6], Si Hammond[1], Andy Herdman[5], Mike Heroux[1], Stephen Jarvis[2], Paul Lin[1], Justin Luitjens[7], Andrew Mallinson[2], Simon McIntosh-Smith[4], Sue Mniszewski[6], Jamal Mohd-Yusof[6], David Richards[3], Christopher Sewell[6], Sriram Swaminarayan[6], Heidi Thornquist[1], Christian Trott[1], Courtenay Vaughan[1], Alan Williams[1], *and your name here*.

1. Sandia National Laboratories
2. University of Warwick, UK
3. Lawrence Livermore National Laboratory
4. University of Bristol, UK
5. AWE, UK
6. Los Alamos National Laboratory
7. Nvidia, Inc.

Pillars of science

Theory

Simulation

Performance models

Architecture simulators

Miniapps on testbeds

Experiment

Application codes are...

- 1M+ SLOC, written using

- MPI + Fortran/C/C++ (+ OpenMP),

- depend on multiple tpls,

- execute on multiple generations of machines from

  multiple vendors,

- capture the work of generations of scientists, and

  are

- getting mission-driven work done today!

Miniapps enable exploration

Focus	Proxy for a key app performance issue
Intent	Tool for codesign: output is information
Scope of change	Any and all
Size	K SLOC
Availability	Open source (LGPL)
Developer/owner	Application team
Life span	Until its no longer useful

Related:

Benchmark	Output: metric to be ranked.
Compact app	Application relevant answer.
Skeleton app	Inter-process comm, application "fake"
Proxy app	Über notion

Notes

- Goal is to tune application, not miniapp.

- Assessing the predictive capabilities of miniapps
  - `http://www.sandia.gov/~rfbarre/PAPERS/Miniapp_vv_SAND.pdf`
    - Methodology informed by experimental V&V
    - Building up body of evidence.
    - Recognition of measurement intrusion.

- Trinity procurement using miniapps:
  - `http://www.nersc.gov/systems/trinity-nersc-8-rfp/draft-nersc-8-trinity-benchmarks/`
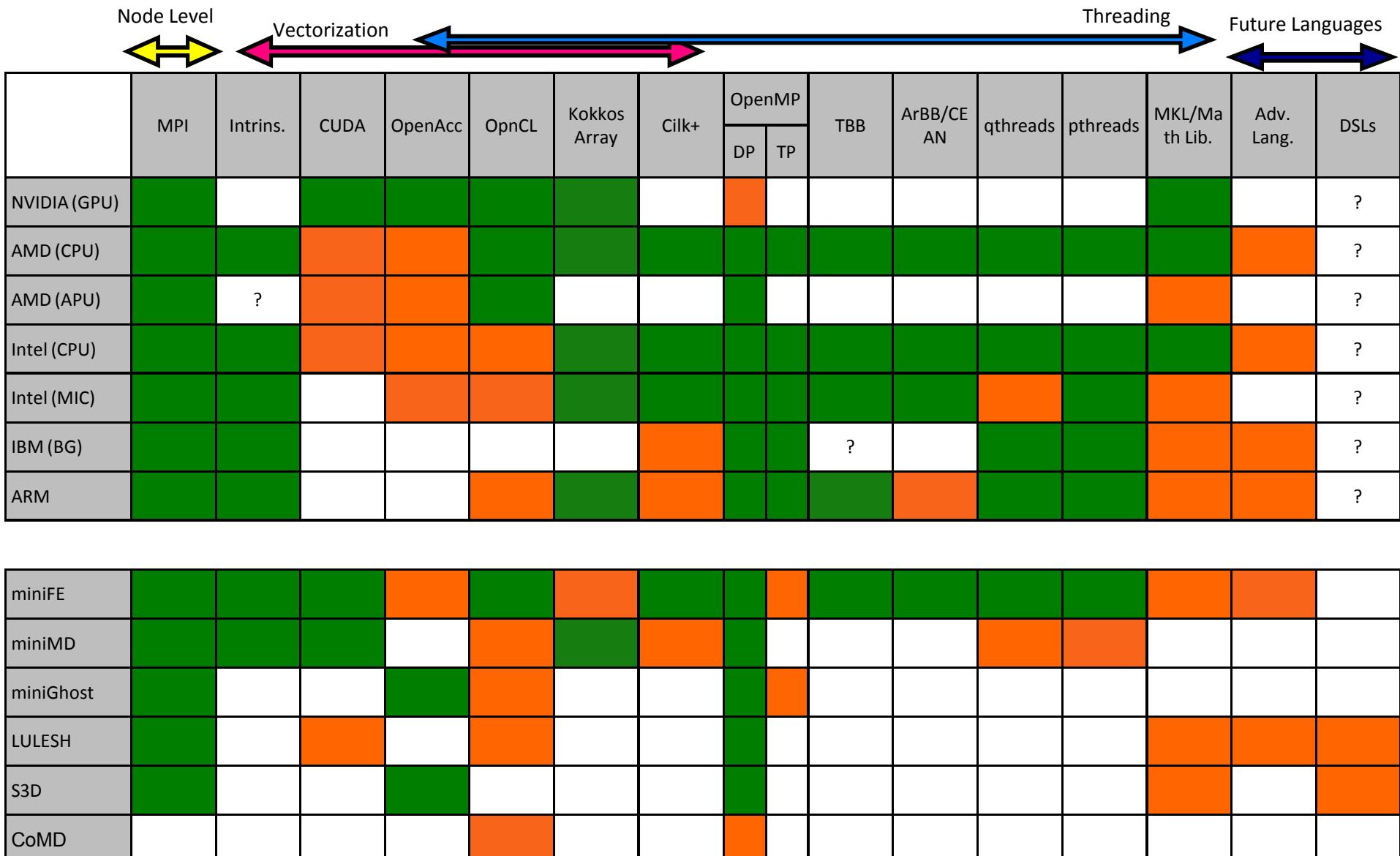
Mantevo project

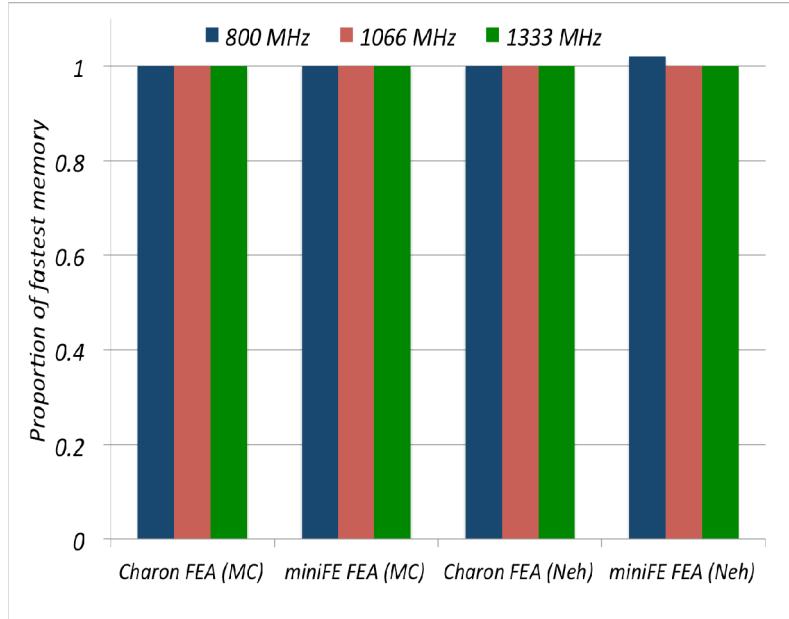| Miniapp	
CloverLeaf/CleverLeaf	Compressible Euler eqns on Cartesian grid, explicit, 2nd order
CoMD	Molecular dynamics, mimics SPaSM.
HPCCG	Conjugate gradient solver
miniFE	Implicit finite element solver
miniGhost/AMR*	FDM/FVM explicit
miniMD	Molecular dynamics (Lennard-Jones)
miniSMAC2D	Finite-differenced 2D incompressible Navier-Stokes
miniXyce	SPICE-style circuit simulator
mini"Aero"**	tbd
miniContact**	Solid mechanics
miniExDyn-FE	Explicit Dynamics (Kokkos-based)
miniITC-FE	Implicit Thermal Conduction (Kokkos-based)
miniWave**	Lagrangian FEM for the hyperbolic wave equation
GraphApp**	Data analytics
PathFinder*	Data analytics

Version 2.0

*in admin
** in development

Snapshot of Technologies

	MPI	Intrins.	CUDA	OpenAcc	OpnCL	Kokkos Array	Cilk+	OpenMP DP	OpenMP TP	TBB	ArBB/CEAN	qthreads	pthreads	MKL/Math Lib.	Adv. Lang.	DSLs
NVIDIA (GPU)	green		green	green	green	green		orange						green		?
AMD (CPU)	green	green	orange	orange	green	green	green	green		green	green	green	green	green	orange	?
AMD (APU)	green	?	orange	orange	green			green						orange		?
Intel (CPU)	green	green	orange	orange	orange	green	green	green		green	green	green	green	green	orange	?
Intel (MIC)	green	green		orange	orange	green	green	green		green	green	orange	green	orange		?
IBM (BG)	green	green					orange	green		?		green	green	orange	orange	?
ARM	green	green			orange	green	orange	green	green	green	orange	green	green	orange	orange	?

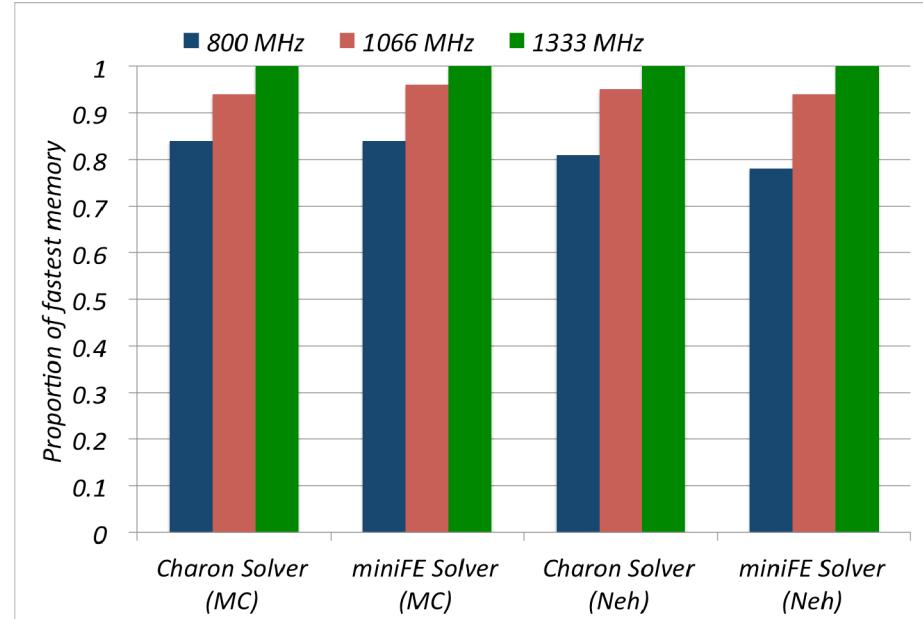|																
miniFE	green	green	green	orange	green	orange	green	green	orange	green	green	green	green	orange	orange	
miniMD	green	green	green		orange	green	orange	green				orange	orange			
miniGhost	green			green	orange			green	orange							
LULESH	green		orange		orange			green						orange	orange	orange
S3D	green			green				green						orange		orange
CoMD					orange			orange								

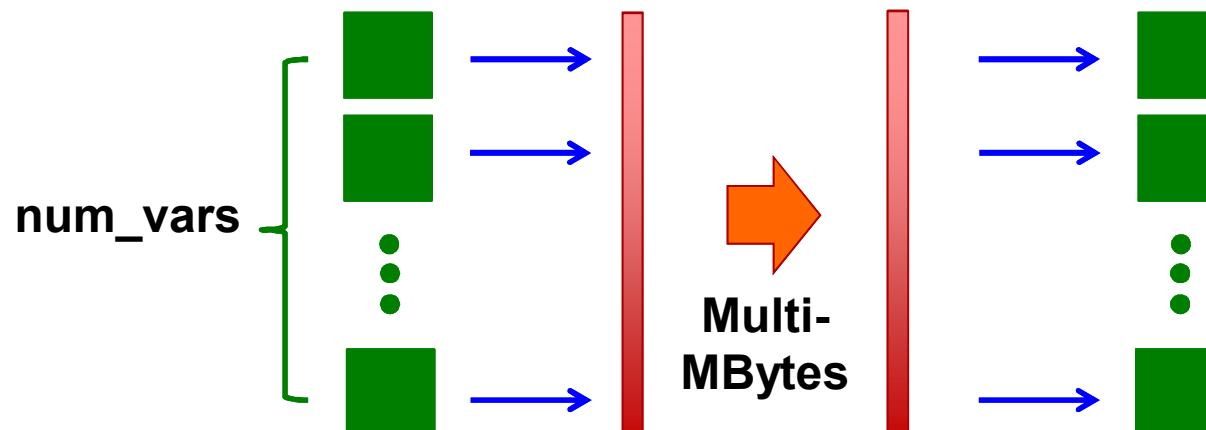# Impact of memory speeds, on Charon and miniFE.

FEA

Solver

MC: AMD Opteron Magny-Cours
Neh: Intel Xeon Nehalem

Revisiting halo exchange strategies

CTH : Eulerian Multi-material modeling application.

- 3-d, finite volume, stencil computation.

- BSP with message aggregation (BSPMA).

- miniGhost Mantevo miniapp configured as proxy.
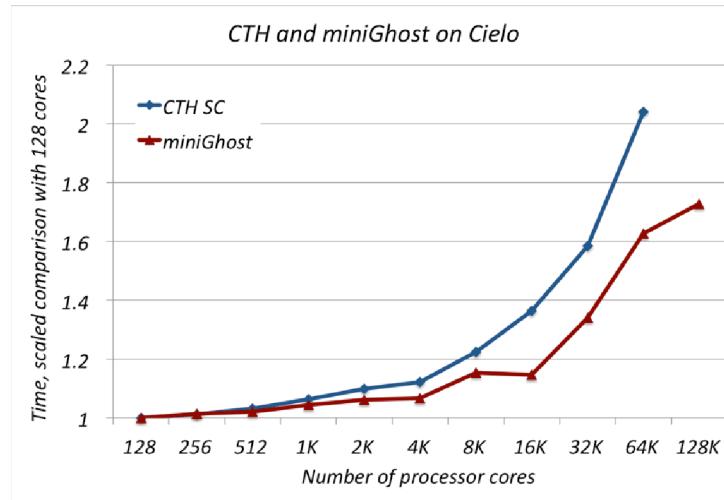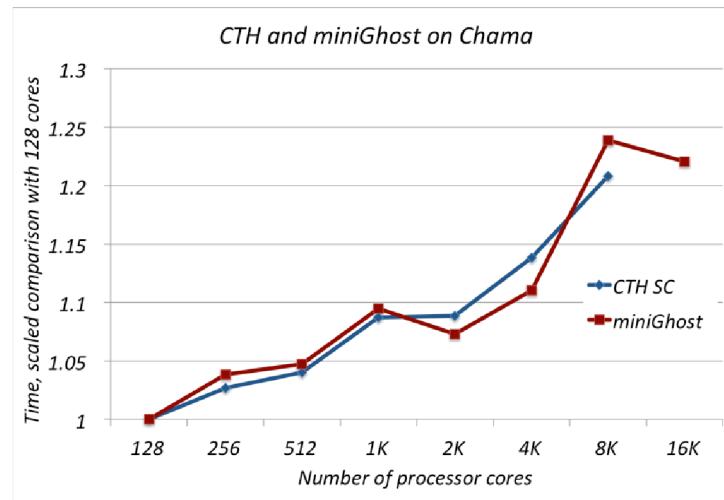
do i = 1, num_tsteps

num_vars

Multi-MBytes

do j = 1, num_vars
  compute
end do
end do

CTH and miniGhost

CTH and miniGhost

Communication Time for MiniGhost (Cielo)

CTH and miniGhost on Cielo, with reordering

CTH and miniGhost on Chama, with reordering
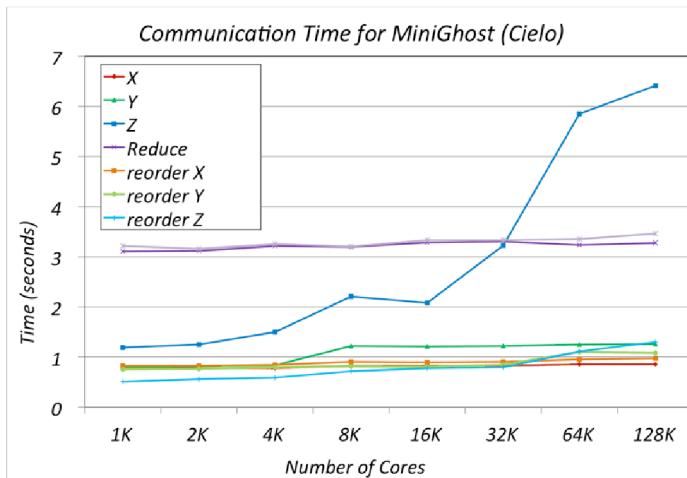
Number of MPI ranks	Regular Order			Reordered		
	X	Y	Z	X	Y	Z
16	0.0	0.0	0.0	0.0	0.0	0.0
32	0.0	0.0	0.0	0.0	0.0	0.0
64	0.0	0.0	0.3	0.0	0.3	0.0
128	0.0	0.0	1.0	0.0	0.5	0.0
256	0.0	0.0	1.0	0.0	0.5	0.3
512	0.0	0.1	2.0	0.0	0.6	0.4
1024	0.0	0.3	2.1	0.2	1.0	0.7
2048	0.0	0.3	2.7	0.3	1.2	1.2
4096	0.0	0.3	3.7	0.3	1.2	1.2
8192	0.0	0.5	5.1	0.2	1.1	2.0
16384	0.0	0.5	4.9	0.2	1.1	2.2
32768	0.0	0.5	5.6	0.2	1.1	2.5
65536	0.0	1.1	10.2	0.2	1.6	2.8
131072	0.0	1.1	10.1	0.2	1.6	3.1
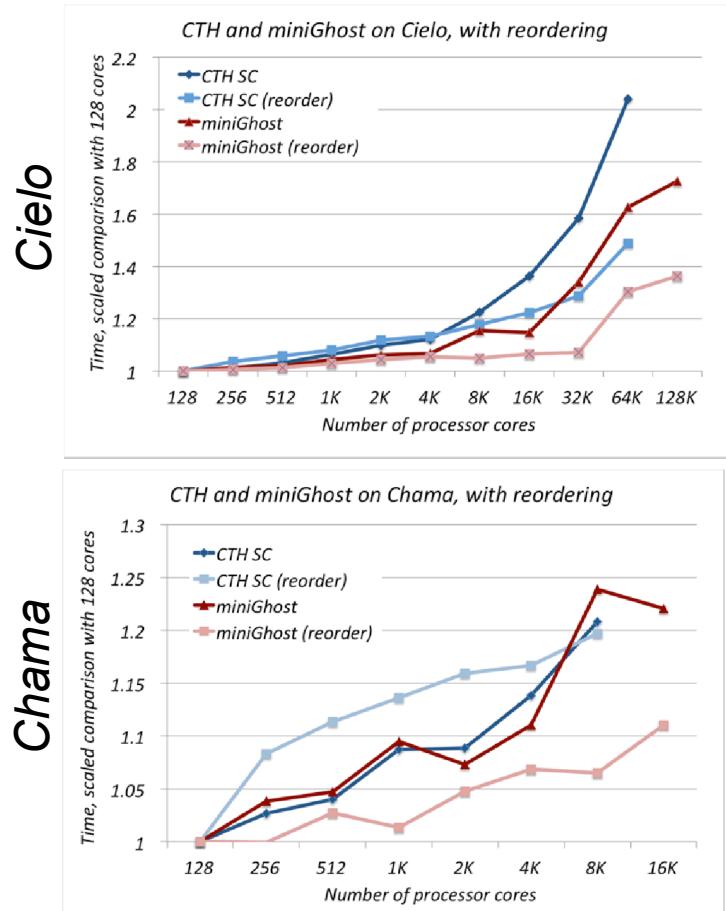
Profiling shows that computation time remains constant, but need more experiments to make stronger claim.

Alternative inter-node strategy : Single Variable Aggregated Faces (SVAF)

Future architectures : less global bandwidth proportional to msg injection rate and bw
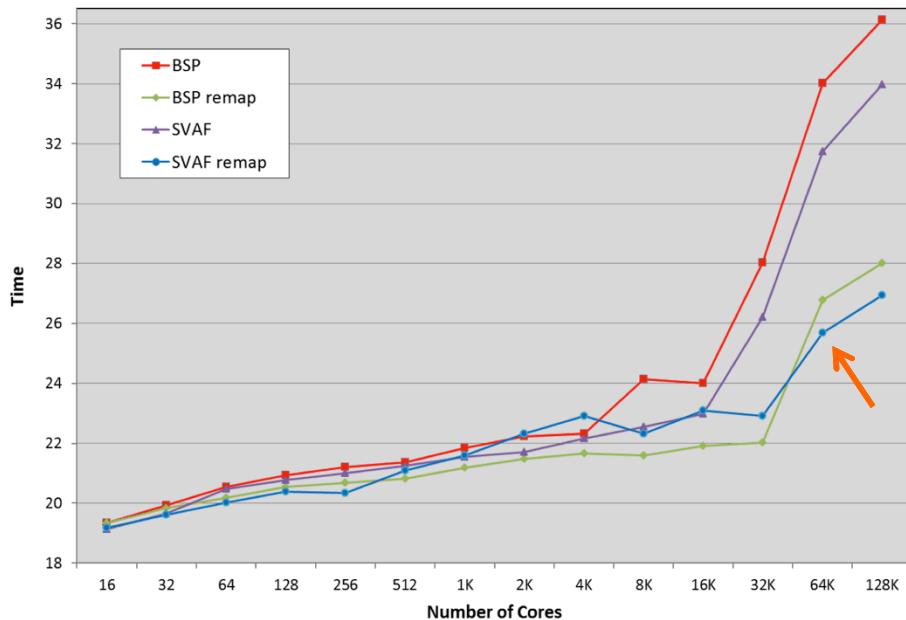
```
do i = 1, num_tsteps

    do j = 1, num_vars
```
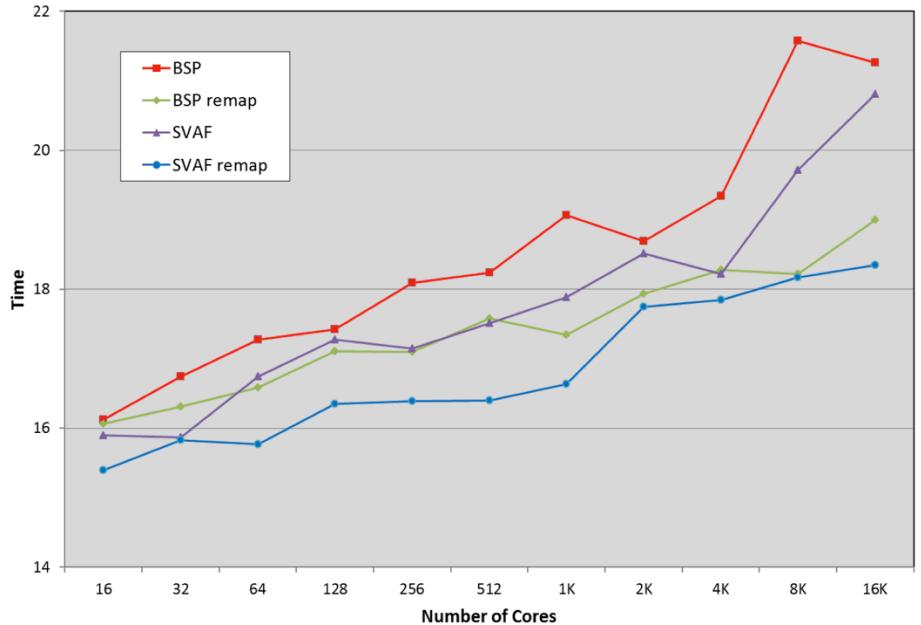

compute

```
    end do

end do
```

SVAF performance

Cielo

Chama

Summary

- Concrete steps can be taken to preparing applications for what we see impacting exascale computation.

- These steps lead to stronger performance on current and emerging architectures.

- Miniapps and testbeds enable a tractable means for exploring these issues.

- Mailing lists: http://mantevo.org/mail_lists.php

On-going work

- Intel ɸ, Kepler, ARM, XC, Tilera, Xeon, Opteron, …

- Revolutionary programming models, languages, mechanisms

Some publications involving Mantevo

- *Statistical Performance Modeling of Modern Out-of-Order Processors using Monte Carlo Methods*, W. Alkohlani , PhD Dissertation, New Mexico State University, 2013.

- *GPU Acceleration of Data Assembly in Finite Element Methods and Its Energy Implications*, Li Tang, Xiaobo Sharon Hu, Danny Ziyi Chen, Michael T. Niemier, Richard F. Barrett, Simon D. Hammond and Ming-Yu Hsieh, The 24th IEEE International Conference on Application-specific Systems, Architectures and Processors (ASAP13), 2013

- *Towards Performance Predictive Application-Dependent Workload Characterization*, W. Alkohlani and J. Cook, In Performance Modeling, Benchmarking and Simulation of High Performance Computer Systems (PMBS12), 2012.

- *Exascale Work-load Characterization and Architecture Implications*, Balaprakash, Buntinas, Chan, Guha, Gupta, S. Narayanan, Chien, Hovland, and Norris. Tech report ANL/MCS-P3013-0712, Argonne National Laboratory, 2012.

- *A Case for Dual Stack Virtualization: Consolidating HPC and Commodity Applications in the Cloud*, Kocoloski, Ouyang, and Lange. In Proceedings of the ACM Symposium on Cloud Computing, SoCC '12, New York, NY, 2012.

- *Report of Experiments and Evidence for ASC L2 Milestone 4467 - Demonstration of a Legacy Application's Path to Exascale*, B.W. Barrett et al. Technical Report, Sandia National Laboratories, 2012.

- *Micro-applications for Communication Data Access Patterns and MPI Datatypes*, Schneider, Gerstenberger, and Hoefler. In Proceedings of EuroMPI, Lecture Notes in Computer Science, volume 7490. Springer, 2012.

- *Exploring Latency-power Tradeoffs in Deep Non-Volatile Memory Hierarchies*, Yoon, Gonzalez, Ranganathan, and Schreiber. In Proceedings of the 9th conference on Computing Frontiers, CF '12, New York, NY, 2012. ACM.

- *Task Mapping for Noncontiguous Allocations*, Feer, Rhodes, Price, Bunde, and Leung. Under review.

- *Communication Optimization Beyond MPI*, Friedley and Lumsdaine. In Proceedings of the 2011 IEEE International Symposium on Parallel and Distributed Processing Workshops and PhD Forum, IPDPSW '11, 2011.

- *On the Viability of Checkpoint Compression for Extreme Scale Fault Tolerance*, Ibtesham, Arnold, Ferreira, and Bridges. In Proceedings of the 2011 International Conference on Parallel Processing - Volume 2, Euro-Par'11.

- *Charm++ for Productivity and Performance*, L. Kale et al., 2011.

- *A performance model with a fixed point for a molecular dynamics kernel*, Numrich and Heroux. Computer Science - Research and Development, 23(3-4), June 2009.

Some publications involving Mantevo

- *Improving Performance via Mini-applications*. Heroux et al. Technical Report, Sandia National Laboratories, 2009.

- *On the Role of Co-design in High Performance Computing*, R.F. Barrett, S. Borkar, S.S. Dosanjh, S.D. Hammond, M.A. Heroux, X.S. Hu, J.Luitjens, S.Parker, J. Shalf, and L.Tang, Transition of HPC Towards Exascale Computing, E.H. D'Hollander et al. (Eds.), IOS Press, 2013.

- *Reducing the Bulk of the Bulk Synchronous Parallel model*, R.F. Barrett, C.T. Vaughan, and S.D. Hammond. Parallel Processing Letters, 2013.

- *Assessing the Validity of the Role of Mini- Applications in Predicting Key Performance Characteristics of Scientific and Engineering Applications*, Barrett, Crozier, Hammond, Heroux, Lin, Trucano, and Vaughan. Under review.

- *Application Explorations for Future Interconnects,* Barrett, Vaughan, Hammond, and Roweth. In Workshop on Large Scale Parallel Processing, at the IEEE International Parallel & Distributed Processing Symposium (IPDPS) Meeting, 2013.

- *Using the Cray Gemini Performance Counters*, Pedretti, Vaughan, Hemmert, and Barrett. Cray User Group Meeting 2013.

- *Emerging High Performance Computing Systems and Next Generation Engineering Analysis Applications*, Ang, Barrett, Hammond, and Rodrigues. In Pacific Rim Workshop on Innovations in Civil Infrastructure Engineering. National Taiwan University of Science and Technology, 2013.

- *Exascale Design Space Exploration and Co-design*. Barrett, D.W. Doerfler, S.S. Dosanjh, Hammond, K.S. Hemmert, Heroux, P.T. Lin, Lutjiens, Pedretti, Rodrigues, and Trucano. Future Generation Computer Systems, special issue on Extreme Scale Parallel Architectures and Systems. To appear

- *Navigating An Evolutionary Fast Path to Exascale*, Barrett, Hammond, Vaughan, D.W. Doerfler, Heroux, Luitjens, and Roweth. In Performance Modeling, Benchmarking and Simulation of High Performance Computer Systems (PMBS12), 2012. Extended version: Technical Report, Sandia National Laboratories, 2012.

- *Summary of Work for ASC L2 Milestone 4465: Characterize the Role of the Mini-Application in Predicting Key Performance Characteristics of Real Applications*, Barrett, Crozier, Hammond, Heroux, Lin, Trucano, and Vaughan. Technical Report, Sandia National Laboratories, 2012.

- *Toward Codesign in High Performance Computing Systems*, Barrett, Hu, Dosanjh, Parker, Heroux, and Shalf. In Proceedings of the International Conference on Computer- Aided Design, ICCAD'12, New York, NY, USA, 2012. ACM.

- *Exascale Computing and the Role of Co-design*. In High Performance Computing: From Grids and Clouds to Exascale, chapter. J. Ang et al. IOS Press Inc, 2011.

- *MiniGhost: A Miniapp for Exploring Boundary Exchange Strategies Using Stencil Computations in Scientific Parallel Computing*, Barrett, Vaughan, and Heroux. Technical Report, Sandia National Laboratories, May 2011.

- *Mantevo views: A flexible system for gathering and analyzing data for the mantevo project*, Christesen. College of St. Benedict/St. John's University Senior Honors Thesis, 2007. Undergraduate Thesis.