



The Performance Effect of Multi-core on Scientific Applications

Jonathan Carter, Yun (Helen) He, John Shalf,
Erich Strohmaier, Hongzhang Shan,
and Harvey Wasserman

NERSC
Lawrence Berkeley National Laboratory

CUG 2007, May 7-10, Seattle, WA



Outline

- **Introduction and Micro-benchmarks**
- **Application Studies**
 - MILC
 - BeamBeam3D
- **Performance Prediction for Multi-core Applications**
 - Model Introduction
 - Model Verification with Various Applications
 - Quad Core Performance Prediction
- **Conclusion**



Current Trend

- **New Constraints**
 - 15 years of *exponential* clock rate growth has ended
- **But Moore's Law continues!**
 - Number of transistors keep increase exponentially.
 - How do keep performance increasing at historical rates?
- **Industry Response**
 - #cores per chip doubles every 18 months *instead* of clock frequency!

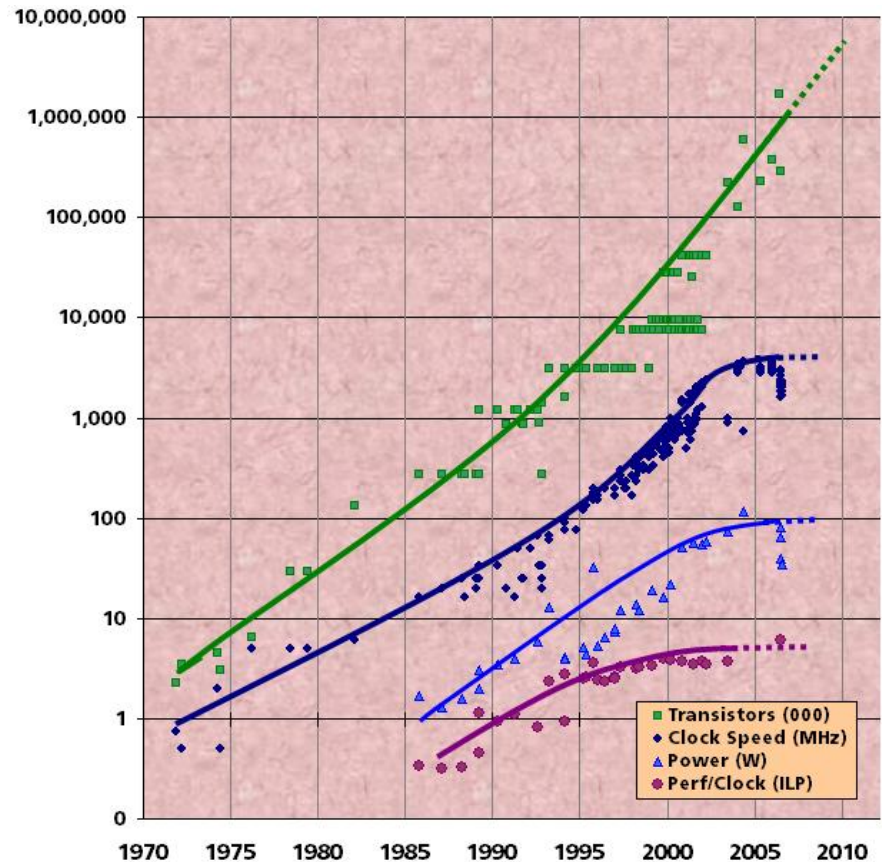


Figure courtesy of Kunle Olukotun, Lance Hammond, Herb Sutter, and Burton Smith



Impact to NERSC

- **Franklin Upgrade Option**
 - Currently 19,000 dual core XT4 2.6GHz Rev-F Opterons
 - Have option to upgrade to quad-core in 2008
 - What is impact of dual-core on application performance
 - Can we use the dual-core impact to predict impact of quad-core on application performance?
 - Ultimately is the quad-core upgrade cost-effective?
- **For Users**
 - What are the causal factors for multi-core performance loss?
 - How to mitigate the dual-core performance impact?
 - What can we learn from micro-benchmarks and some typical scientific applications?

CUG 2007, May 7-10, Seattle, WA



Previous Work

Understanding and Mitigating Multicore Performance Issues on the AMD Opteron Architecture

John Levesque, Jeff Larkin, Martyn Foster, Joe Glenski, Garry Geissler
Cray Inc

Brian Waldecker
AMD Inc.

Jonathan Carter, David Skinner, Helen He, John Shalf, Harvey Wasserman
LBNL/NERSC

Hongzhang Shan, Erich Strohmaier
LBNL/CRD

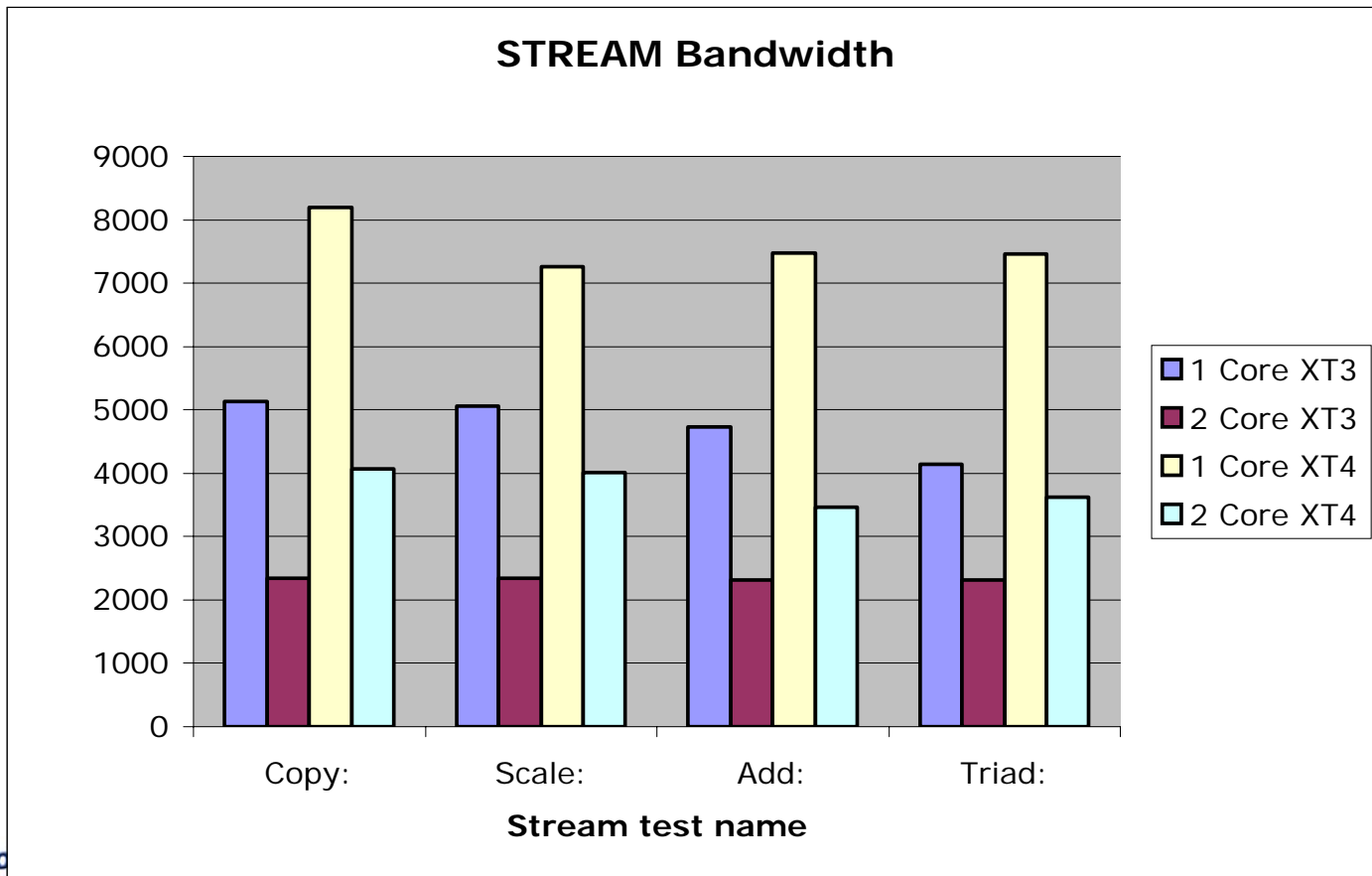
LBNL-62500, March 2007

CUG 2007, May 7-10, Seattle, WA



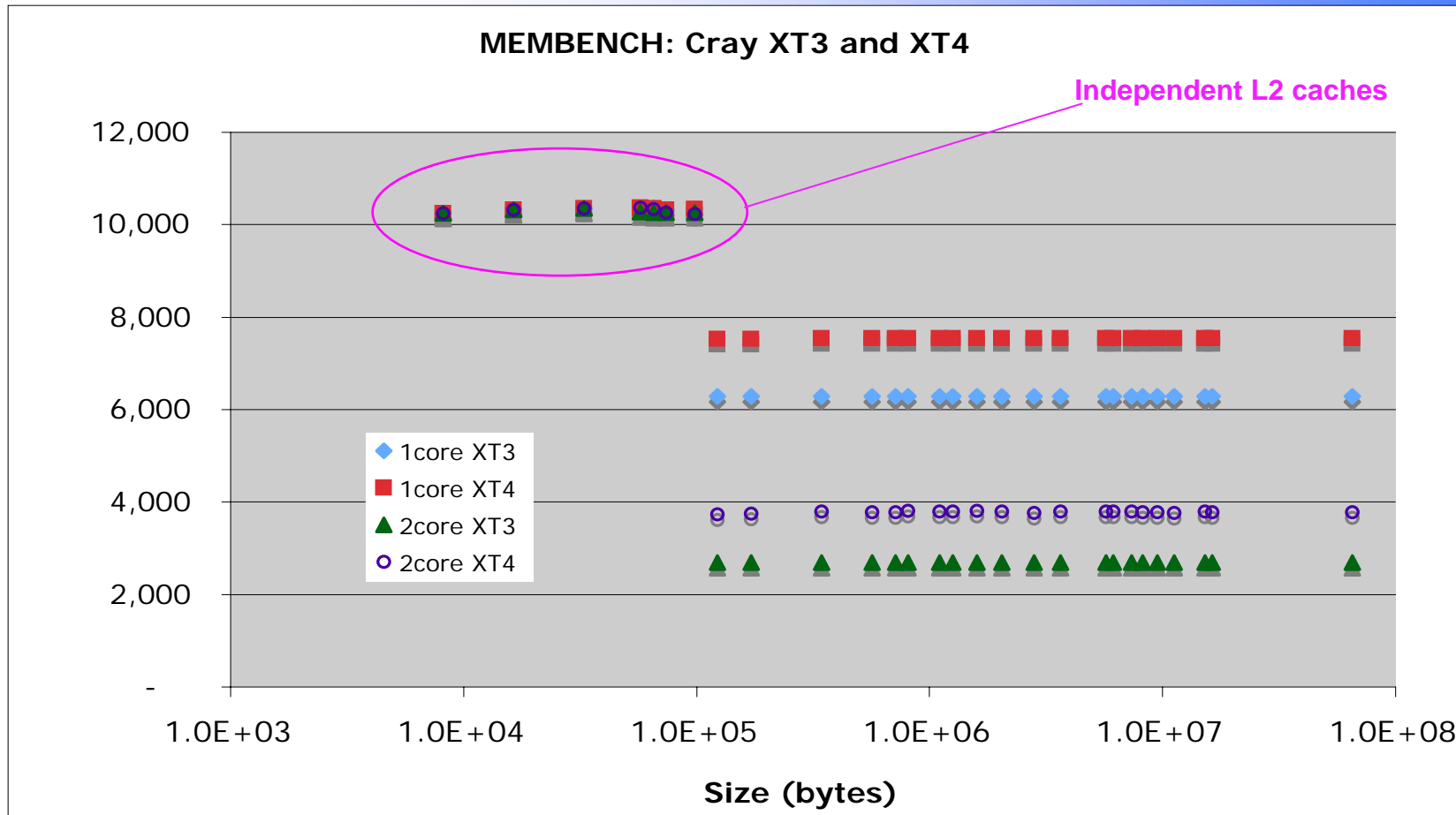
STREAM

	1 Core XT3	1 Core XT4	2 Core XT3	2 Core XT4
Copy:	5137	8196	2345	4074
Scale:	5067	7257	2348	4012
Add:	4734	7482	2309	3469
Triad:	4135	7464	2310	3626





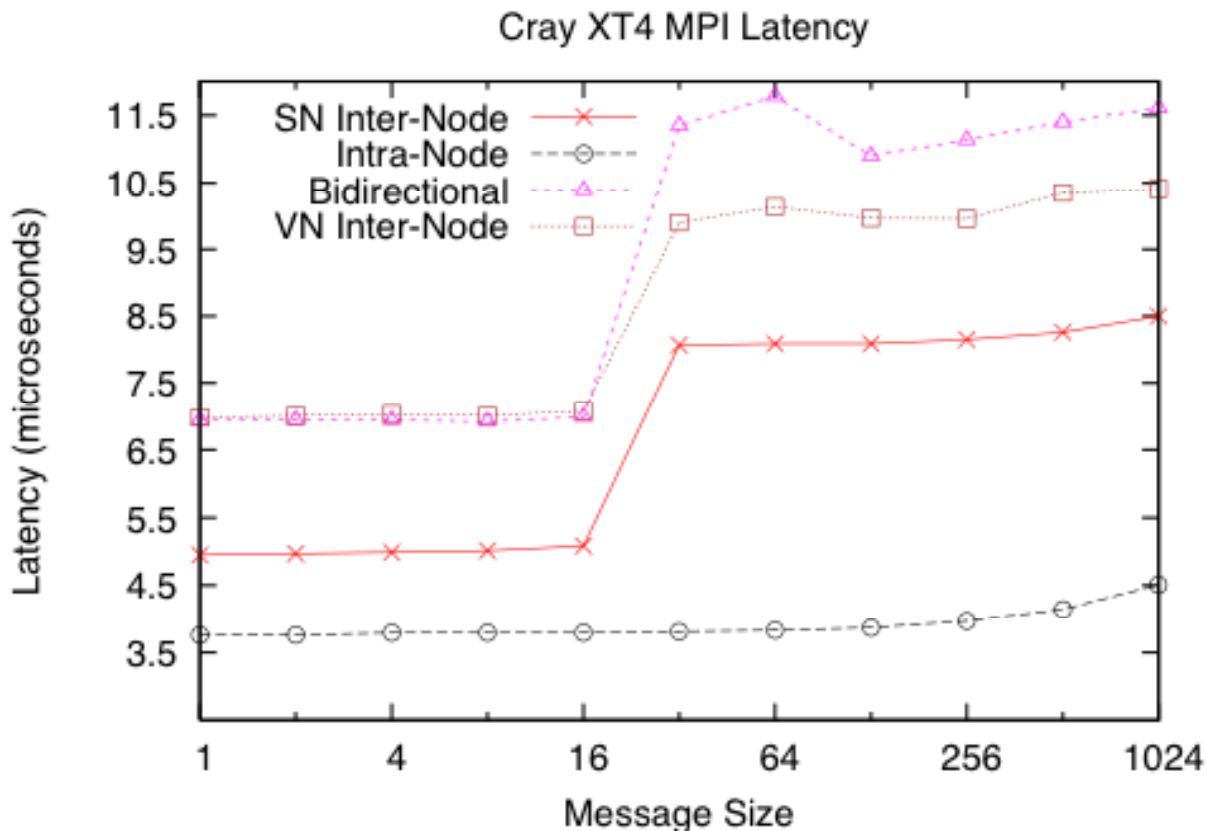
Membench Memory Bandwidth



ftn -tp k8-64 -fastsse -Minfo -Mnontemporal Mprefetch=distance:8,nta

CUG 2007, May 7-10, Seattle, WA

MPI Latency

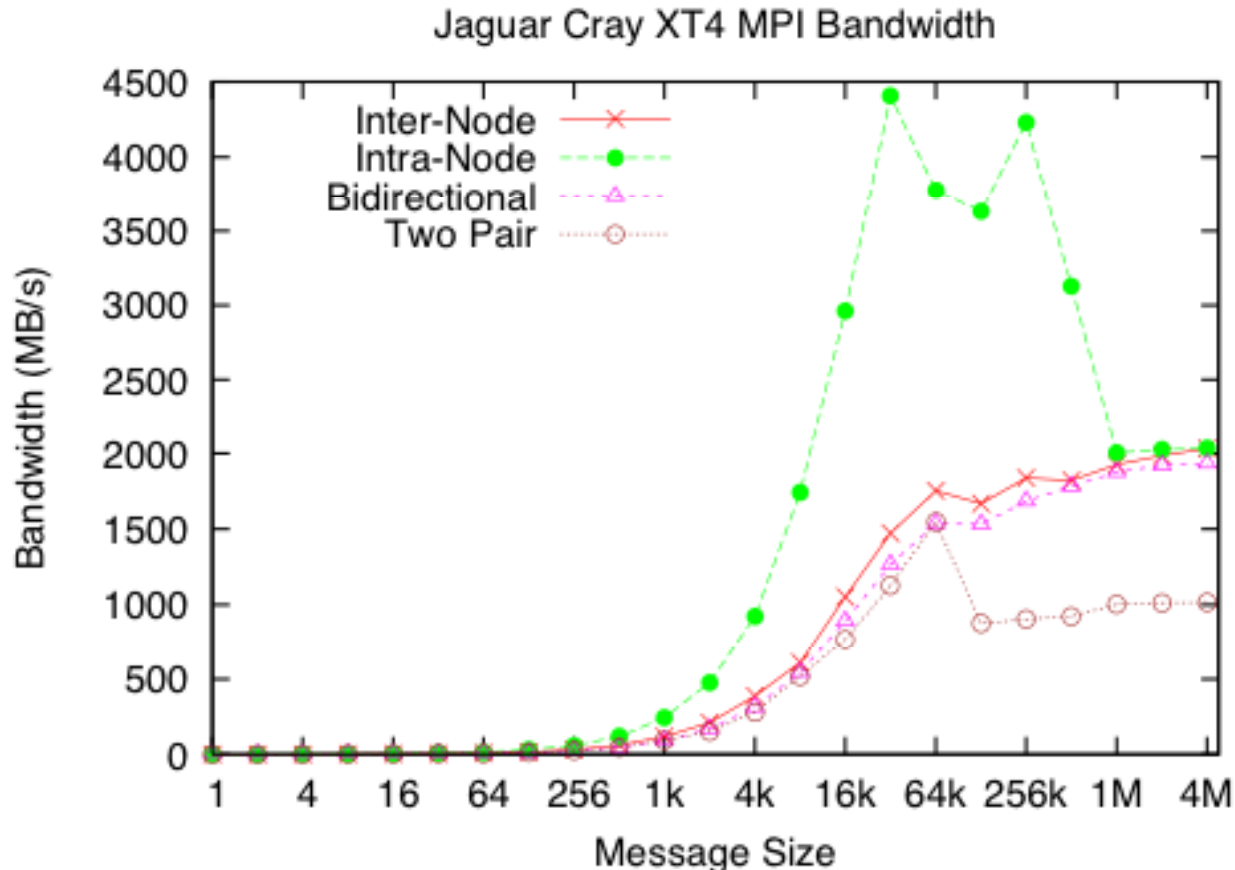


- **MPI latency measured with zero-size message on Jaguar:**
 - Single core inter-node 4.8 usec
 - Dual core inter-node 6.3 usec

CUG 2007, May 7-10, Seattle, WA



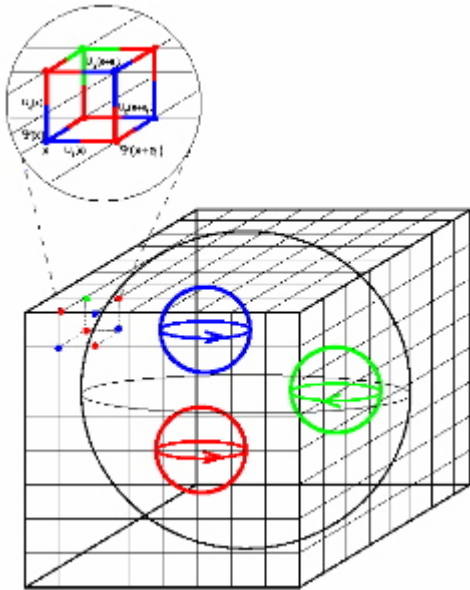
MPI Message Bandwidth



- **Effective MPI bandwidth drops to about half the rate from within a node to between two nodes with 64k message size (typical for MILC)**

CUG 2007, May 7-10, Seattle, WA

MIMD QCD: MILC



A proton on the lattice,
Courtesy www.usqcd.org

- MIMD Lattice Quantum ChromoDynamics (QCD) application
- Widespread community use
 - Easy to build, no dependencies, standards conforming
 - Can be setup to run on wide-range of concurrency
- Conjugate gradient algorithm
- Physics on a 4D lattice
- Local computations are 3x3 complex matrix multiplies, with sparse (indirect) access pattern

MILC on Jaguar

64 cores

	Small Pages		Large Pages	
	Single	Dual	Single	Dual
XT3				
Wall Clock Time	160	230	166	232
Sustained MFLOPS	69370	48402	67138	47976
Percent of Peak	21%	15%	20%	14%
Computational Intensity	2.1	2.1	2.1	2.1
OPS/TLB Miss	308	309	68	68
OPS/D1 Cache Miss	16	16	16	16
OPS/L2 Cache Miss	32	32	31	31
XT4				
Wall Clock Time	127	181	130	184
Sustained MFLOPS	87840	61482	85447	60538
Percent of Peak	26%	18%	26%	18%
Computational Intensity	2.1	2.1	2.1	2.1
OPS/TLB Miss	307	308	106	106
OPS/D1 Cache Miss	16	16	16	16
OPS/L2 Cache Miss	33	33	33	33

- Problem: 32^4 lattice with two trajectories of five steps each.
- SSE inlined assembly with aggressive prefetching.
- Oddly, relatively little data reuse but still high computational intensity.



MILC Dual Core Penalty

MILC Version	Times (sec)	
	XT3	XT4
Single Core Orig	274	230
Single Core Opt	160	127
Dual Core Orig	358	277
Dual Core Opt	230	181
	Dual Core Penalty	
	XT3	XT4
Original	1.31	1.20
Optimized	1.44	1.43

- > 40% dual core penalty for optimized version.
- Un-optimized version shows lower dual-core penalty.
- Optimization to make better use of memory bandwidth results in greater dual-core penalty.

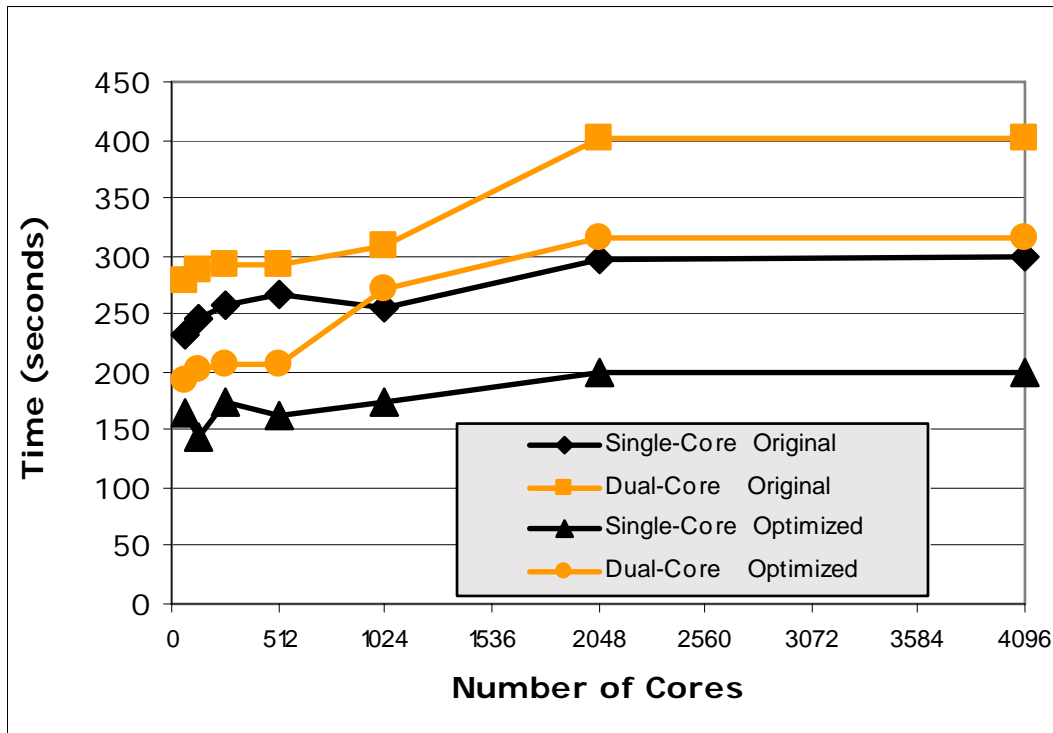


MILC XT4/XT3 Improvement

MILC Version	Improvement: XT4/XT3	
Single Core Orig	1.19	
Single Core Opt	1.26	
Dual Core Orig	1.29	
Dual Core Opt	1.27	
	Improvement: Optimized / Original	
	XT3	XT4
Single Core	1.71	1.81
Dual Core	1.56	1.53

- XT4/XT3 improvement high except for single core un-optimized version.
- Single task of un-optimized version could not saturate the XT4 memory interface, thus not gaining full benefit of improved XT4 memory bandwidth.

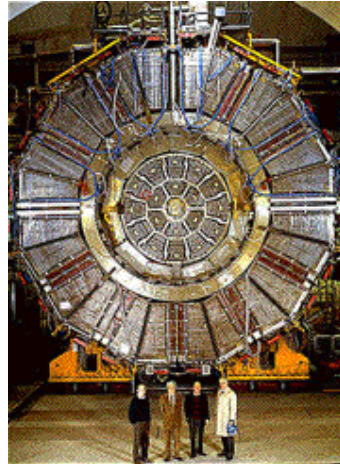
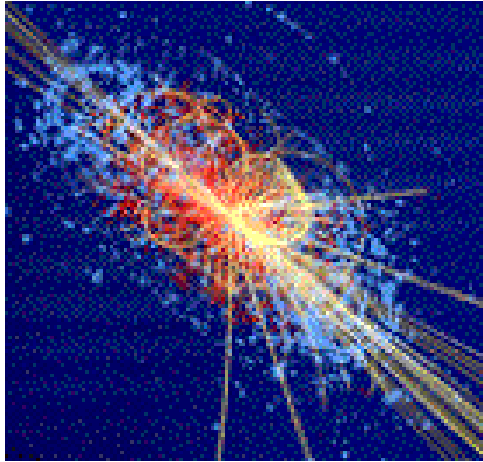
Jaguar XT4



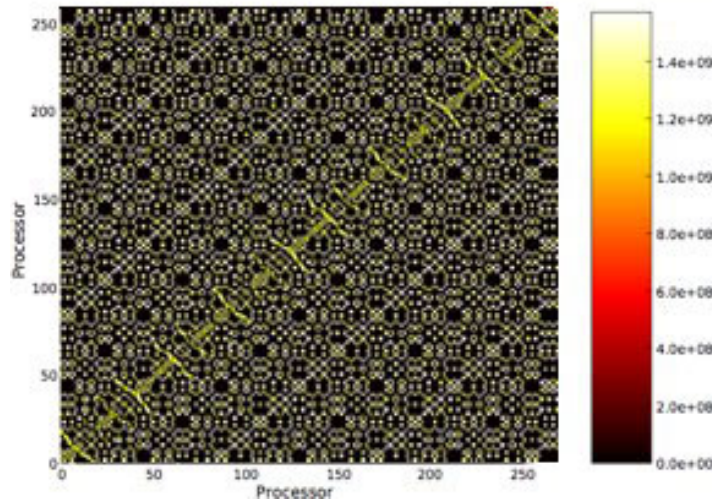
- Un-optimized version with single core runs faster than optimized version with dual core for 1024+ cores.
- Dual core penalty higher with optimized version.
 - Un-optimized version
 - 20%, 64 cores
 - 35%, 4096 cores
 - Optimized version
 - 40%, 64cores
 - 58%, 4906 cores



High Energy Physics: BeamBeam3D



- BB3D models beam-beam collisions of counter-rotating charge particle beams
- Particle -in-cell method, where particles are deposited on 3D grid to calculate charge density distribution
- At collision points electric/magnetic fields calculated using Vlasov-Poisson via FFT
- High communication requirements:
 - Global gather charge density
 - Broadcast electric/magnetic fields
 - Global FFT transpose



CUG 2007, May 7-10, Seattle, WA

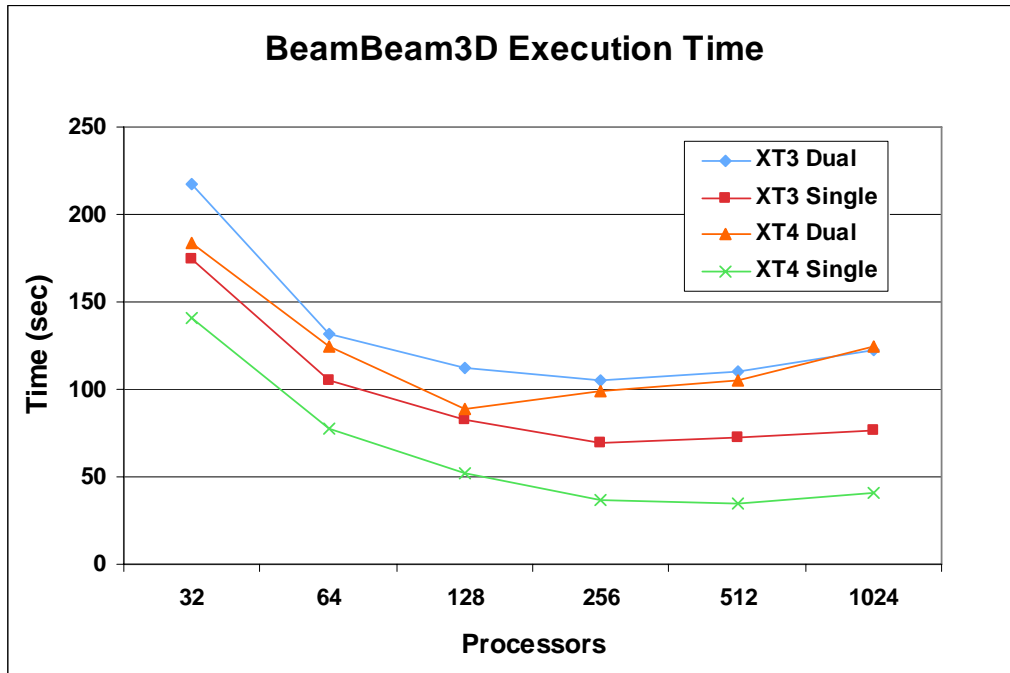


BeamBeam3D on Jaguar

64 cores

Cores	Times (sec)	
	XT3	XT4
Single Core	86	77
Dual Core	109	102
	Dual Core Penalty	
	XT3	XT4
	1.27	1.32

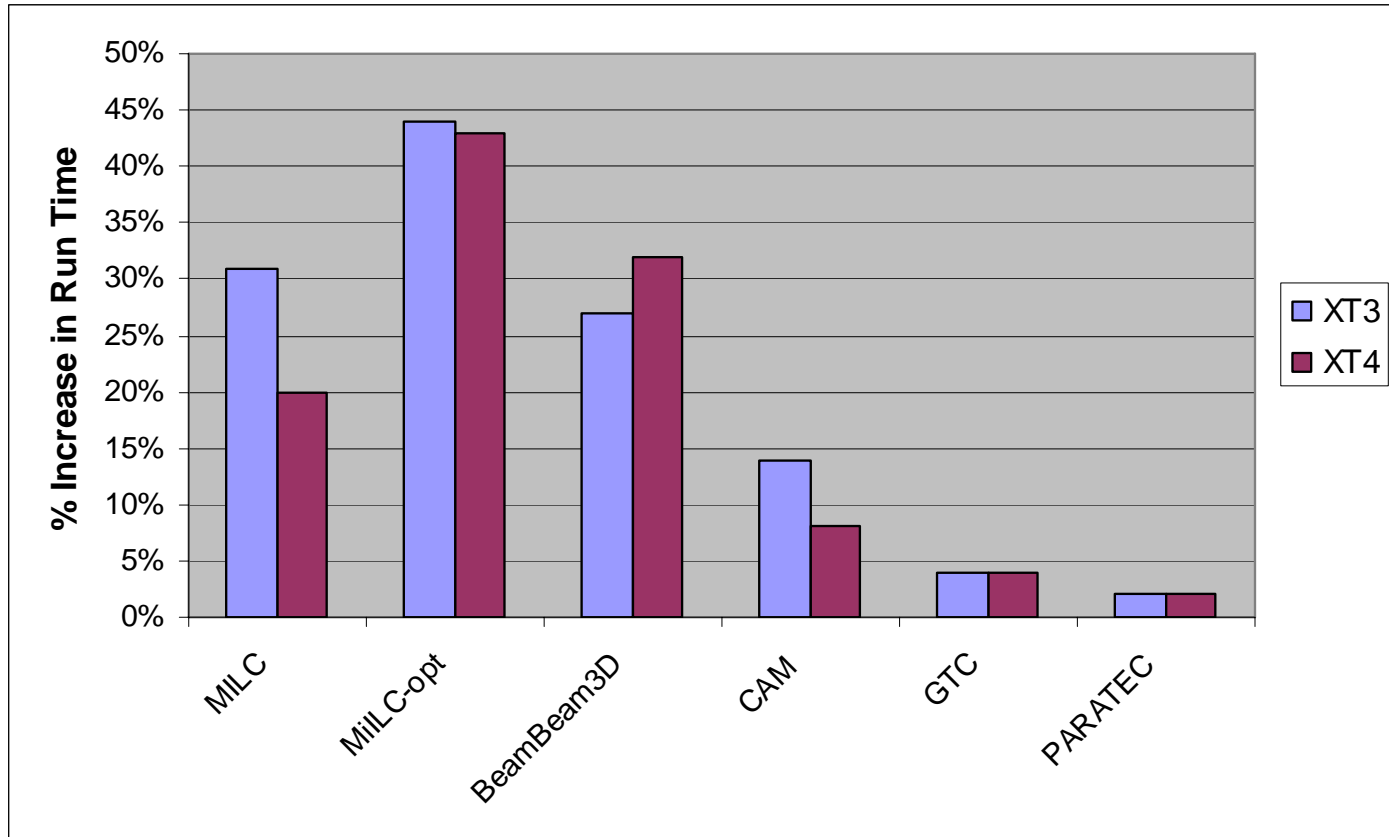
- Problem: 5 million particle simulation with grid resolutions of 256x256x32
- Dual core penalty
 - XT3: 27%
 - XT4: 32%
- XT4/XT3 improvement
 - Single core: 1.12
 - Dual core: 1.07



- **Best Performance on Jaguar**
 - Single core XT3: 256 cores
 - Dual core XT3: 256 cores
 - Single core XT4: 512 cores
 - Dual core XT4: 128 cores
- **Different balance between interconnect and computation in dual-core mode for XT4 node**
 - Large load imbalance
 - Large communication increase at > 128 cores.
 - Major impact on scalability.



Dual Core Performance Penalty



- **MILC, Milc-opt, and BeamBeam3D have higher dual core penalty on Jaguar.**
 - **Memory intensive codes.**

CUG 2007, May 7-10, Seattle, WA



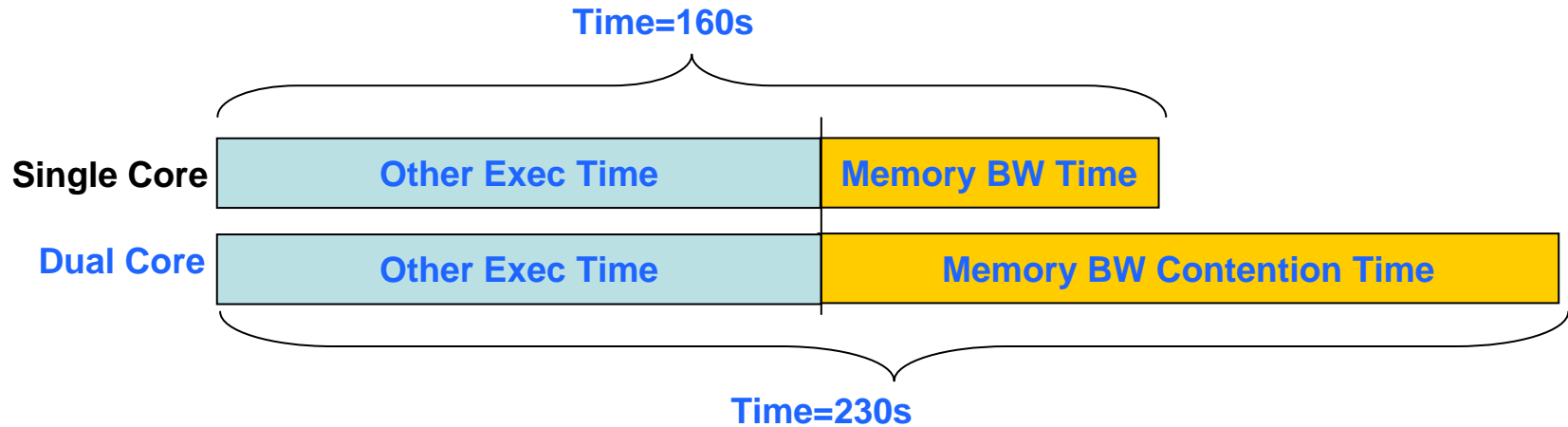
Performance Prediction

- **Model Assumptions**
 - Memory bandwidth is the only contended resource
 - Can break down execution time into portion that is stalled on shared resources (*memory bandwidth*) and portion that is stalled on non-shared resources (*everything else*)
 - Execution time spent using non-shared resources is fixed
 - Estimate time spent on memory contention from XT3 single/dual core studies
 - Estimate # bytes moved in memory-contended zone
 - Extrapolate to XT4 based on increased memory bandwidth
 - Use to validate model
 - Extrapolate to quad-core

CUG 2007, May 7-10, Seattle, WA



Performance Prediction Model

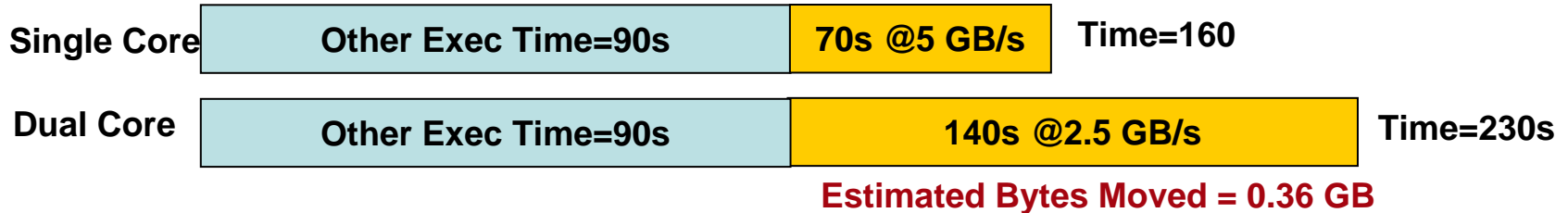


Use MILC-opt XT3 time to illustrate the model

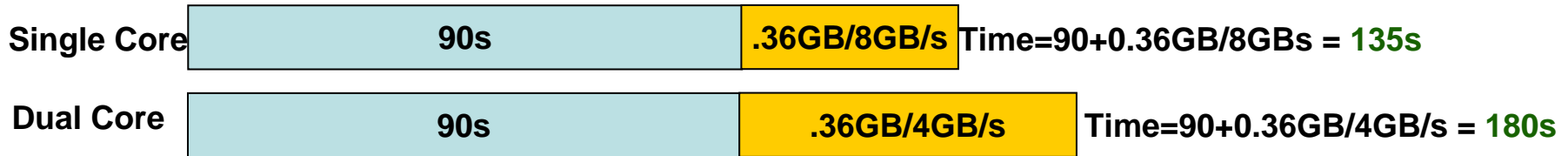


Performance Prediction Model

Cray XT3 Opteron@2.6Ghz DDR400



Cray XT4 Opteron@2.6Ghz DDR2-667



Using actual STREAMS bandwidth data:

MILC-opt Prediction for XT4 SC=120s

actual = 127s, error = -5.1%

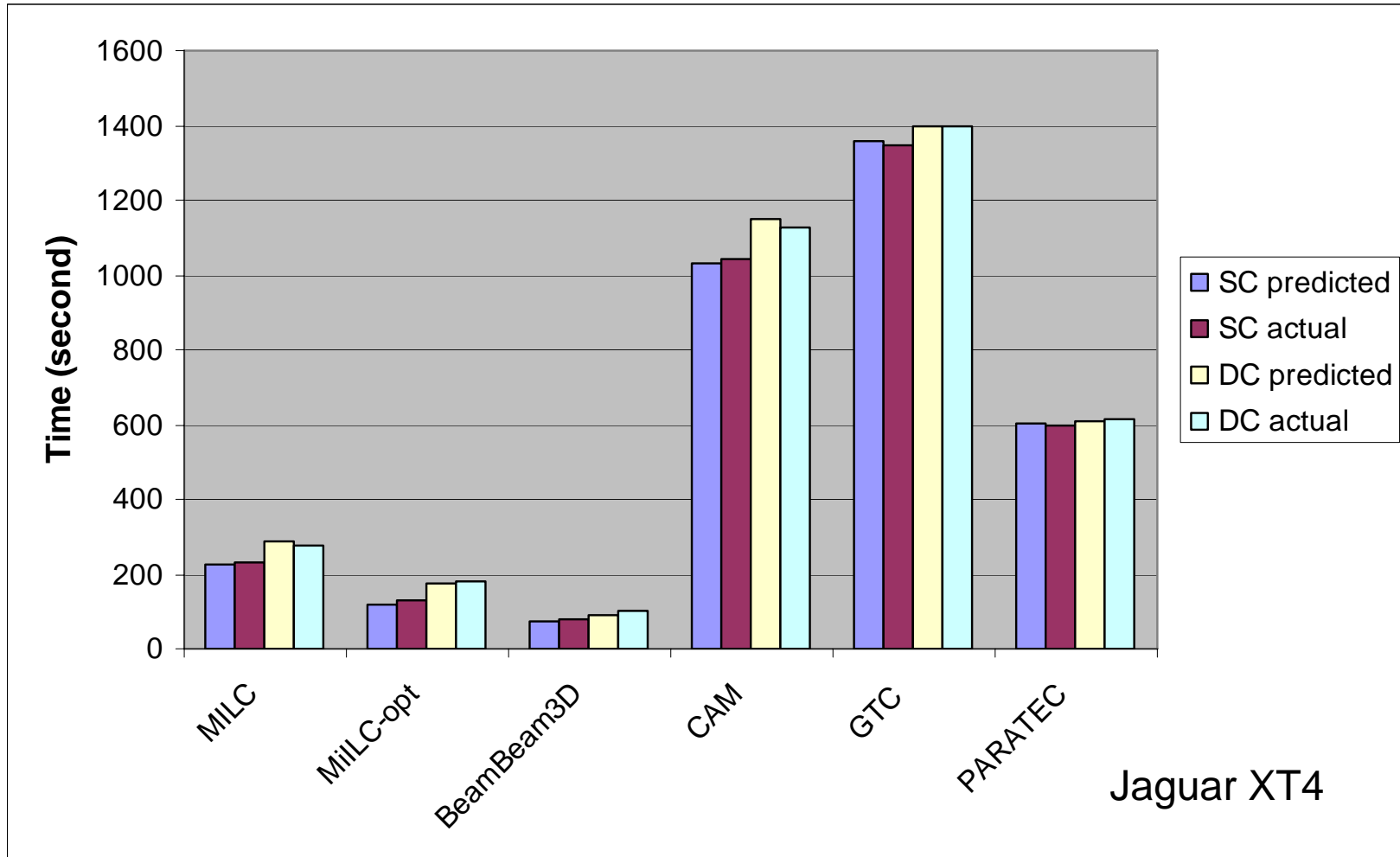
MILC-opt Prediction for XT4 DC = 172s

actual = 181s, error = -4.7%

CUG 2007, May 7-10, Seattle, WA



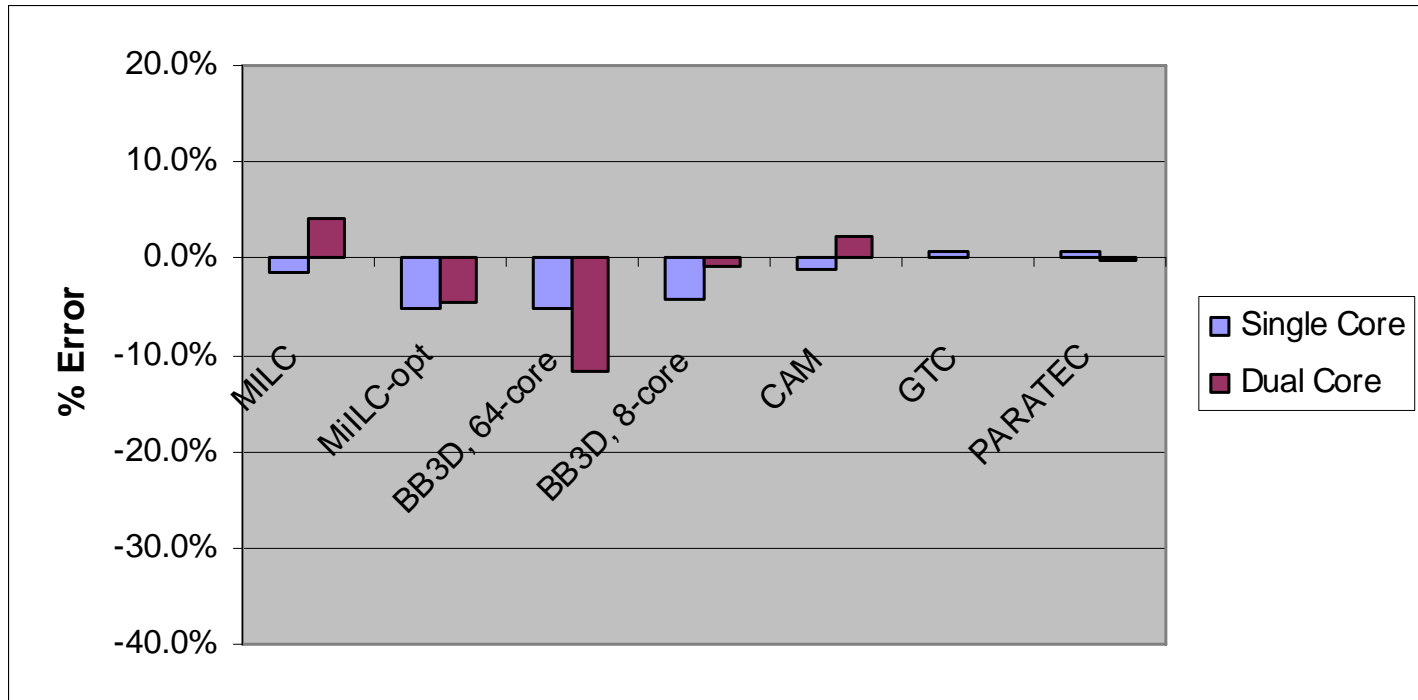
Predicted and Actual Time



CUG 2007, May 7-10, Seattle, WA



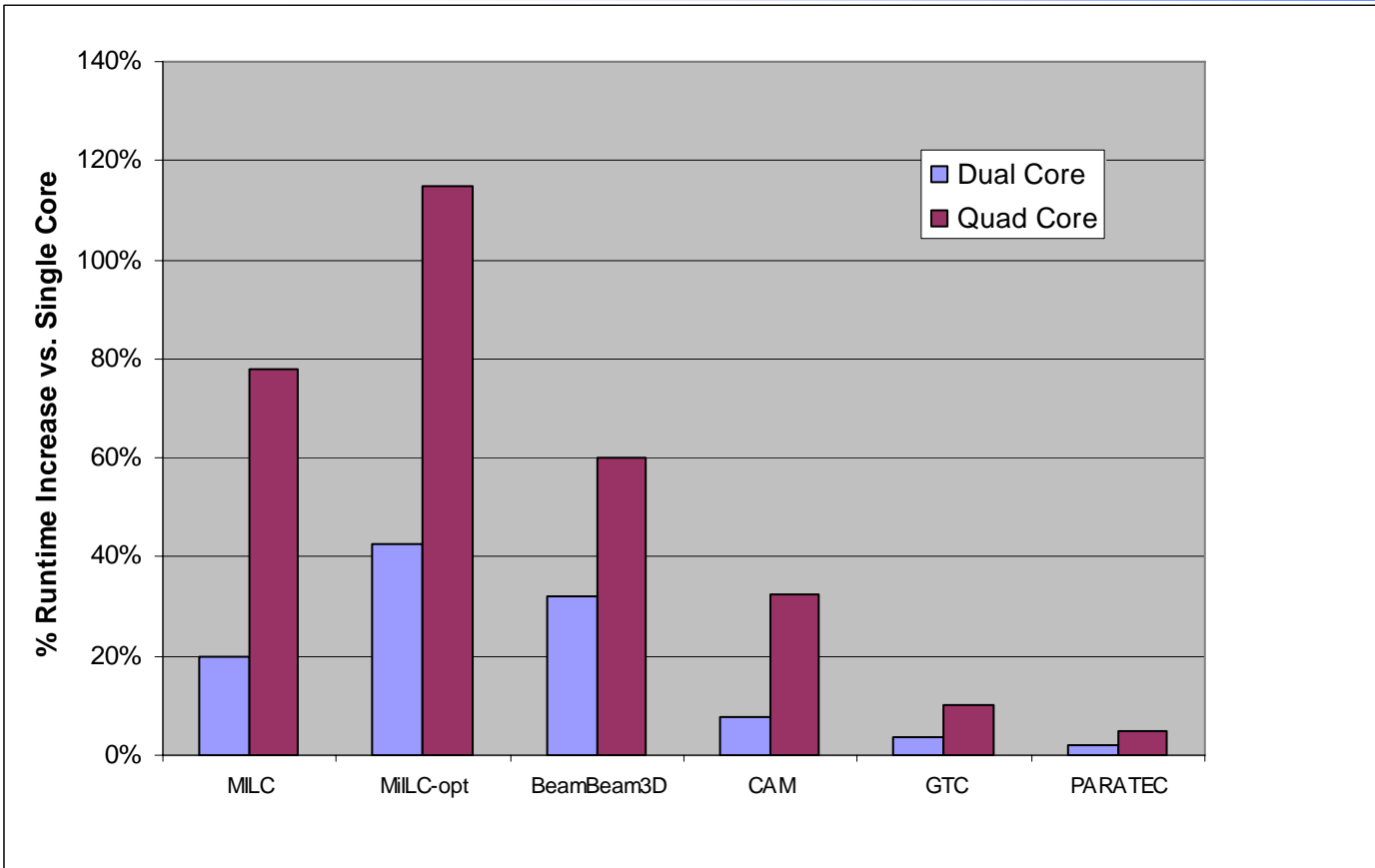
Performance Prediction Error



- Prediction accuracy better than 10%, except for one case.
- Relatively large prediction errors for MILC, MILC-opt, and BeamBeam3D 64-core on Jaguar XT4.
 - Communication effects not accounted for in model
 - Smaller error with BB3D 8-core

CUG 2007, May 7-10, Seattle, WA

Quad Core Prediction



- Quad core penalty large if dual core penalty large.

Conclusions

- **Scaling studies with single and dual core performance of MILC and BeamBeam3D on Jaguar XT3 and XT4. Dual core penalty increases with higher concurrency.**
- **Interesting story from MILC optimization. The aggressive optimization increases memory efficiency, and causes larger dual core penalty.**
- **Performance prediction model introduced. Accuracy verified with various applications using single core and dual core. Model is then used for quad core predictions.**
- **Disclaimer: Quad core prediction**
 - **Assumes no memory bandwidth improvement over dual core.**
 - **Ignores changes to internal cache structures of Opteron.**
 - **Does not take into account the micro-architectural improvements for floating point operations.**



Acknowledgement

- **Technical discussions with Cray and AMD staff**
 - Cray: John Levesque, Jeff Larkin, Martyn Foster, Joe Glenski, Garry Geissler, Stephen Whalen
 - AMD: Brian Waldecker
- **National Center for Computational Sciences at Oak Ridge National Laboratory, supported by the Office of Science, U.S. Department of Energy.**
 - All data collected from Jaguar, most after the recent XT3 and XT4 merge completed on March 26, 2007.
- **Authors supported by the Director, Office of Science, Advanced Scientific Computing Research, U.S. Department of Energy.**

CUG 2007, May 7-10, Seattle, WA