Rethinking Computer Architecture for Throughput Computing

Wen-mei Hwu
University of Illinois, Urbana-Champaign

MulticoreWare, Inc.

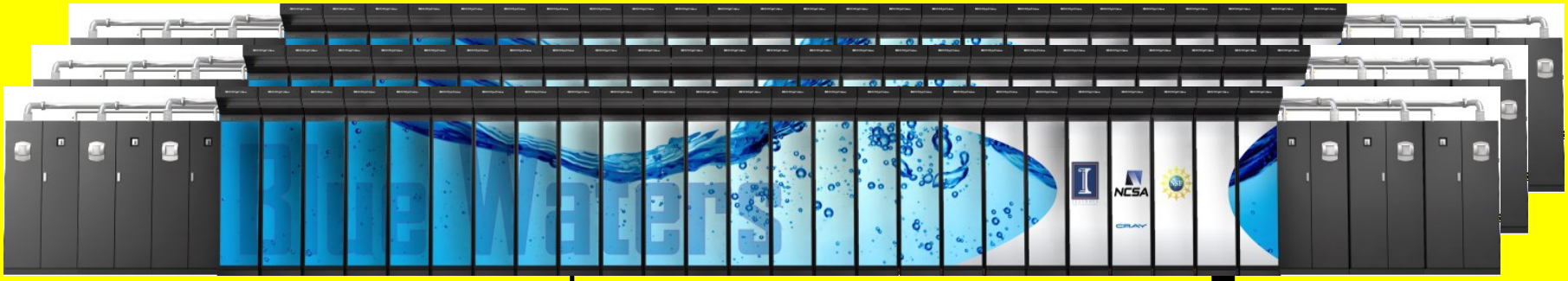
A Few Thoughts on Computer Architecture While in Samos

Agenda

- Where are we today?
- Computer Architecture as the Interface Between Software and Hardware
- Optimizing Across Interfaces
- Computer Architecture as an Academic Discipline

Blue Waters Computing System

Operational at Illinois since 11/2012



11.1 PF
1.5 PB DRAM

10/40/100 Gb Ethernet Switch

IB Switch

>1 TB/sec

120+ Gb/sec

100 GB/sec

WAN

Spectra Logic: 300 PBs

Sonexion: 26 PBs

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Blue Waters and Titan Computing Systems

System Attribute	NCSA Blue Waters	ORNL Titan
Vendors	Cray/AMD/NVIDIA	Cray/AMD/NVIDIA
Processors	Interlagos/Kepler	Interlagos/Kepler
Total Peak Performance (PF)	11.1	27.1
Total Peak Performance (CPU/GPU)	7.1/4	2.6/24.5
Number of CPU Chips	48,352	18,688
Number of GPU Chips	3,072	18,688
Amount of CPU Memory (TB)	1511	584
Interconnect	3D Torus	3D Torus
Amount of On-line Disk Storage (PB)	26	13.6
Sustained Disk Transfer (TB/sec)	>1	0.4-0.7
Amount of Archival Storage	300	15-30
Sustained Tape Transfer (GB/sec)	100	7

Science Area	Number of Teams	Codes	Struct Grids	Unstruct Grids	Dense Matrix	Sparse Matrix	N-Body	Monte Carlo	FFT	PIC	Sig I/O
Climate and Weather	3	CESM, GCRM, CM1/WRF, HOMME	X	X		X		X			X
Plasmas/ Magnetosphere	2	H3D(M),VPIC, OSIRIS, Magtail/UPIC	X				X		X		X
Stellar Atmospheres and Supernovae	5	PPM, MAESTRO, CASTRO, SEDONA, ChaNGa, MS-FLUKSS	X			X	X	X		X	X
Cosmology	2	Enzo, pGADGET	X			X	X				
Combustion/ Turbulence	2	PSDNS, DISTUF	X						X		
General Relativity	2	Cactus, Harm3D, LazEV	X			X					
Molecular Dynamics	4	AMBER, Gromacs, NAMD, LAMMPS				X	X		X		
Quantum Chemistry	2	SIAL, GAMESS, NWChem			X	X	X	X			X
Material Science	3	NEMOS, OMEN, GW, QMCPACK			X	X	X	X			
Earthquakes/ Seismology	2	AWP-ODC, HERCULES, PLSQR, SPECFEM3D	X	X			X				X
Quantum Chromo Dynamics	1	Chroma, MILC, USQCD	X		X	X					
Social Networks	1	EPISIMDEMICS									
Evolution	1	Eve									
Engineering/System of Systems	1	GRIPS,Revisit						X			
Computer Science	1			X	X	X			X		X

So, we can just build bigger and bigger machines, right?

- Not really.
- Blue Waters is running at more than 10M Watt, close to the economic limit for scientific computing
- Technology advancement alone will unlikely produce big breakthroughs
- Weak scaling alone will unlikely result in big discoveries
  - E.g., Molecular Dynamics simulations has been running at about 1 microsecond per day whereas meaning reactions take milliseconds or more

COMPUTER ARCHITECTURE AS THE INTERFACE

Problem

Algorithm

Program

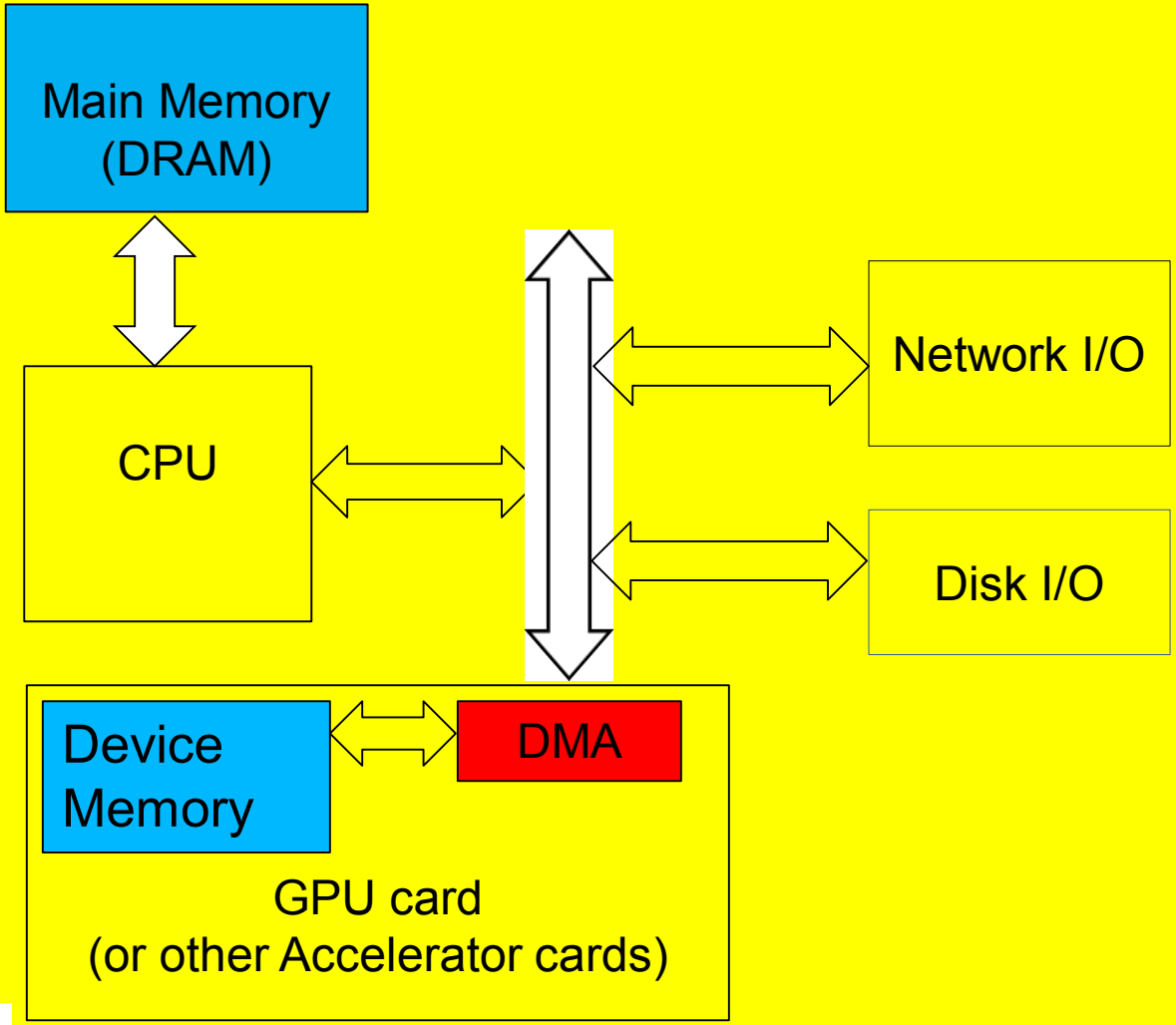
ISA (Instruction Set Arch)

Microarchitecture

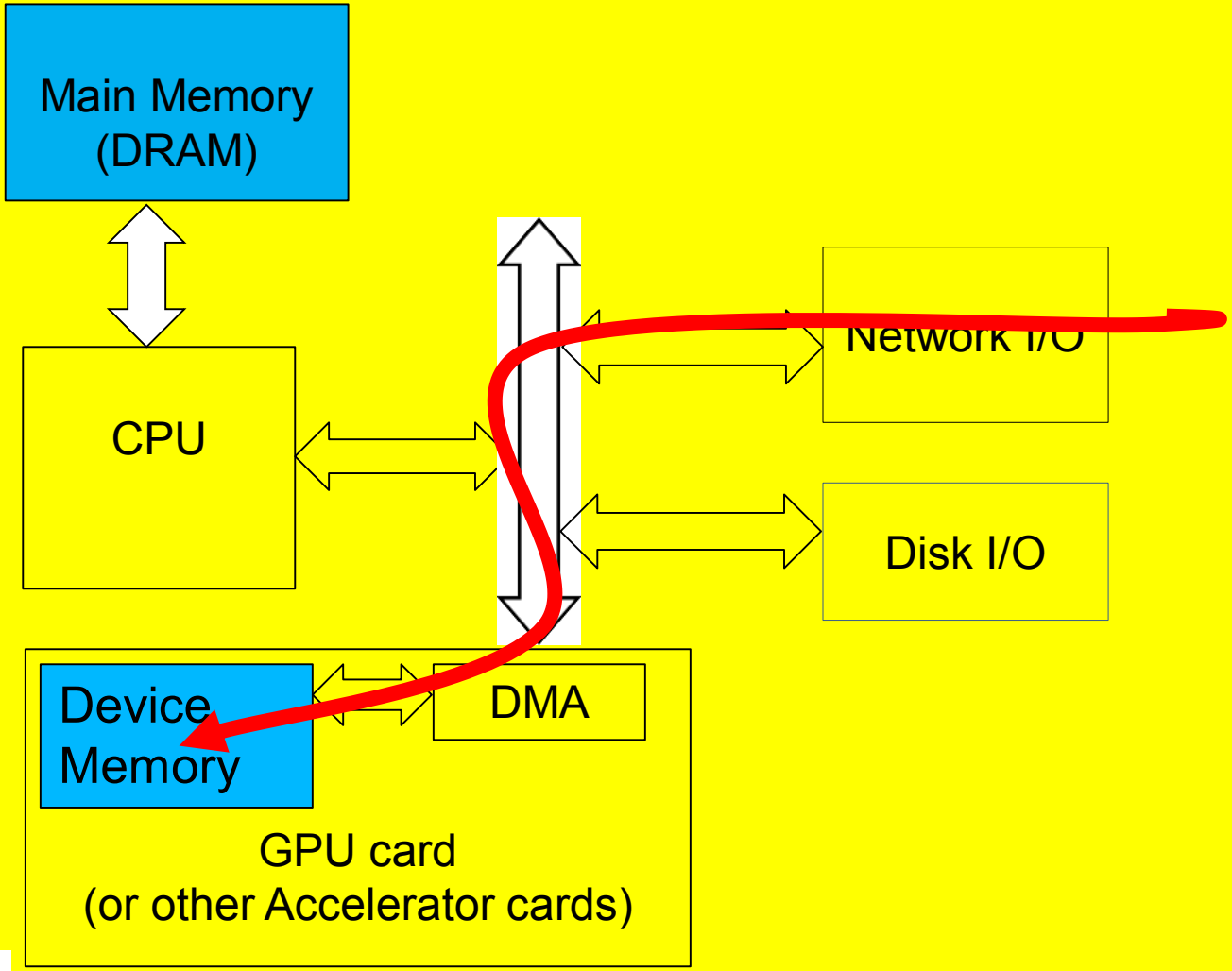
Circuits

Electrons

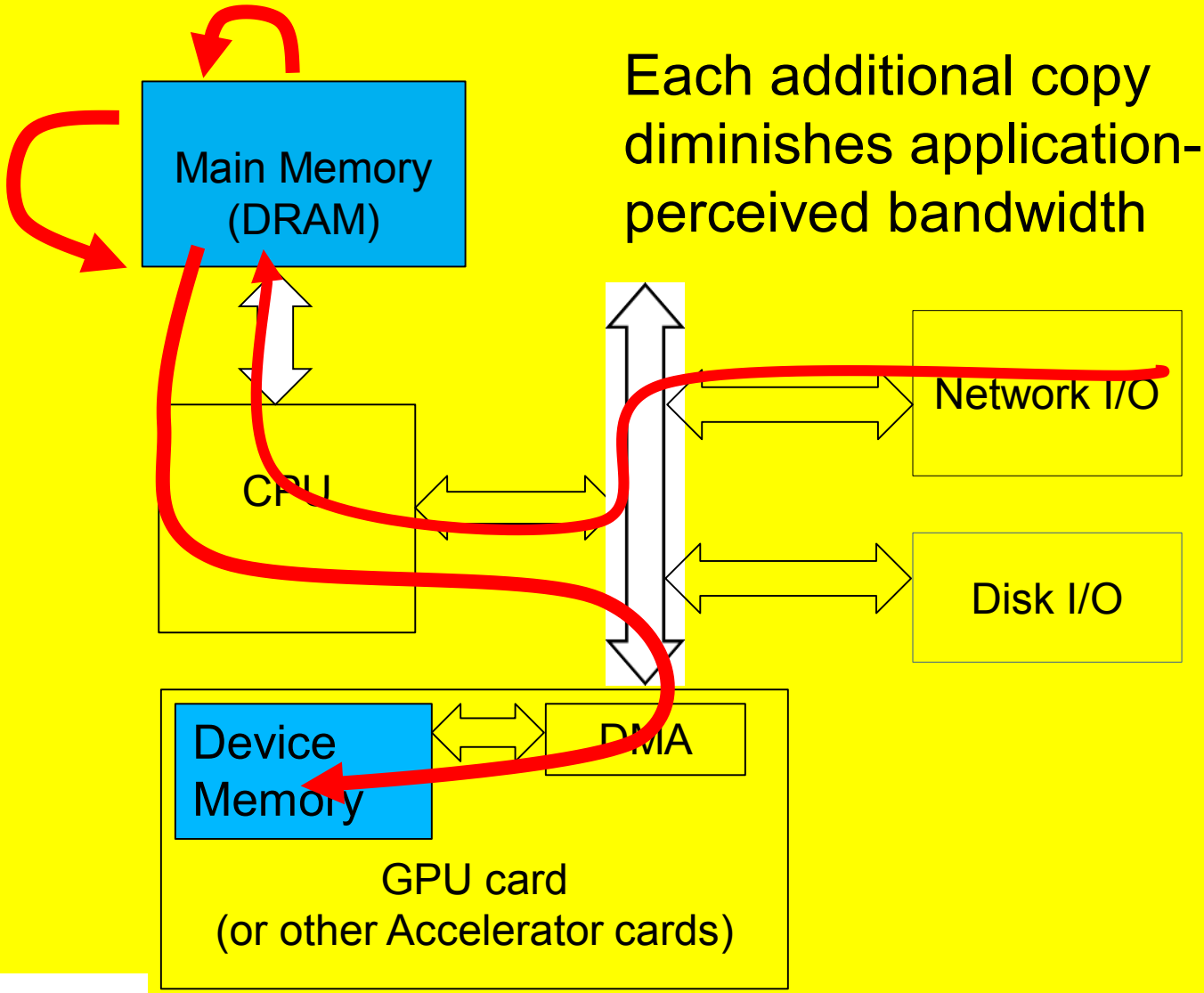
PCIe Data Transfer using DMA



Desirable Data Transfer Behavior



Actual Data Transfer Behavior

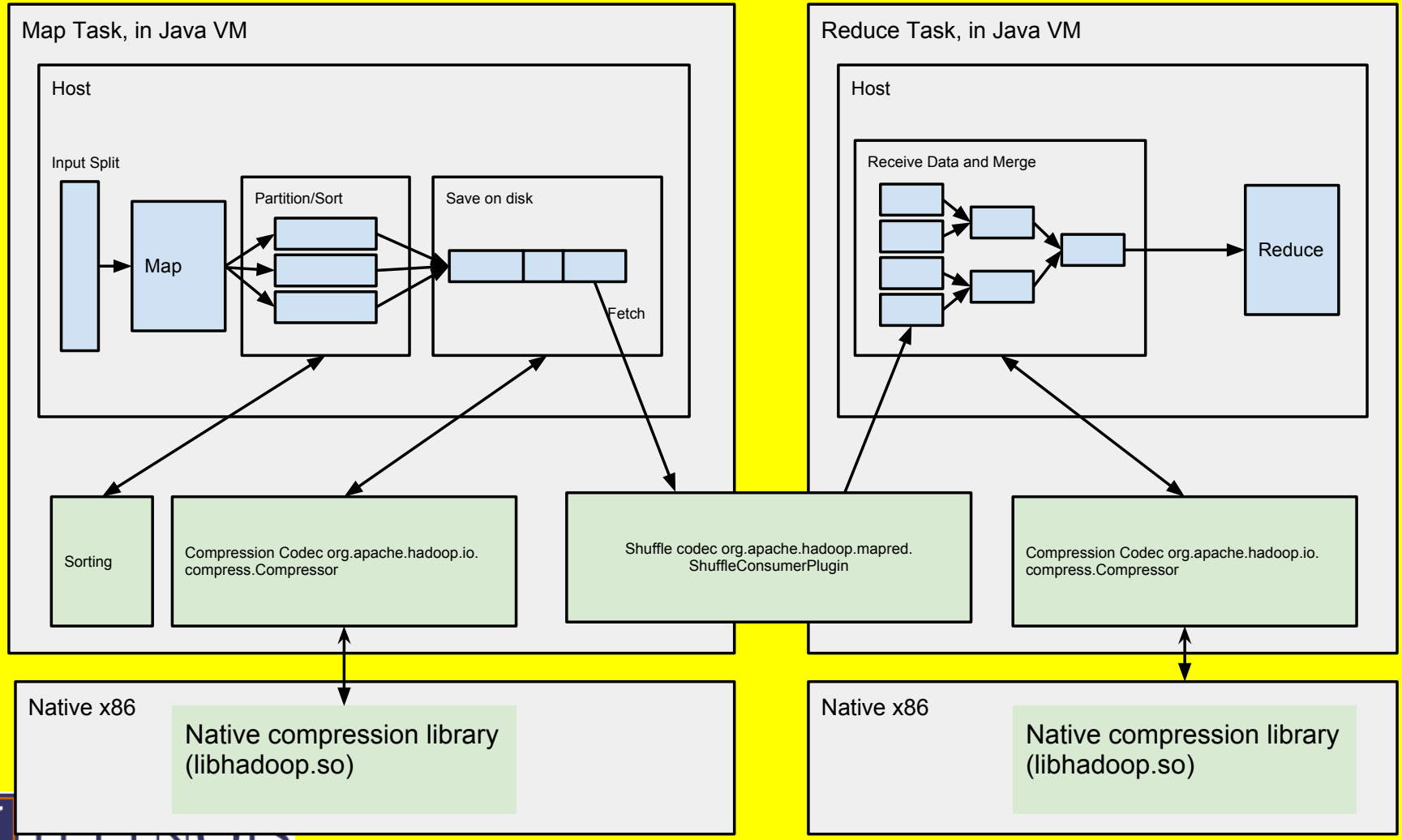


Each additional copy diminishes application-perceived bandwidth

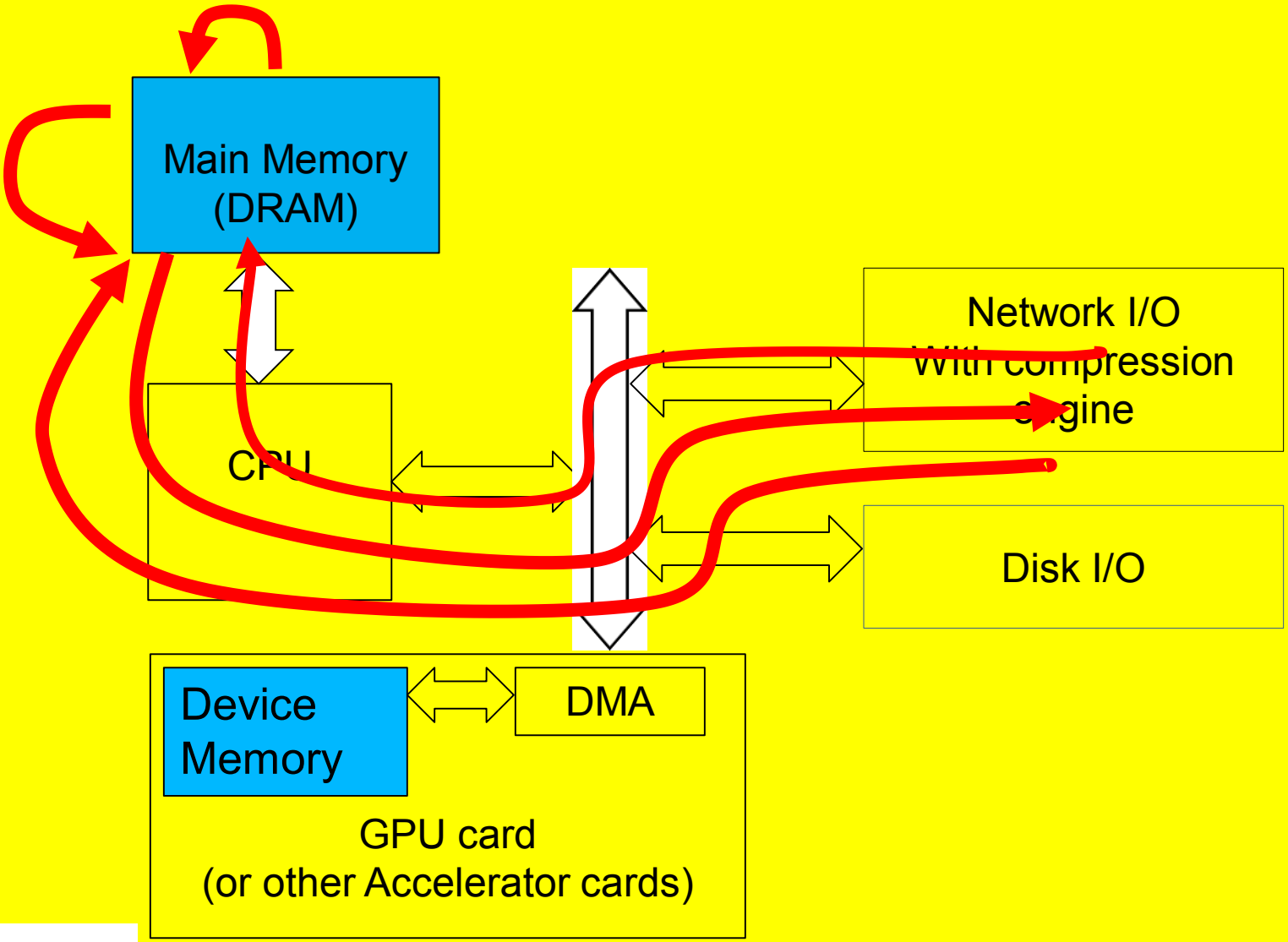
Application Interface Breaks Down

- Older hardware and runtime requires explicit data copy API calls for each copy step
- Newer hardware and runtime allows fewer copies but requires application rewrite
  - Unified physical address space, unified virtual address space, peer to peer DMA
- GMAC (ASPLOS 10) interface abstracts all intermediate copies away
  - Inserts intermediate copies for older HW and RT

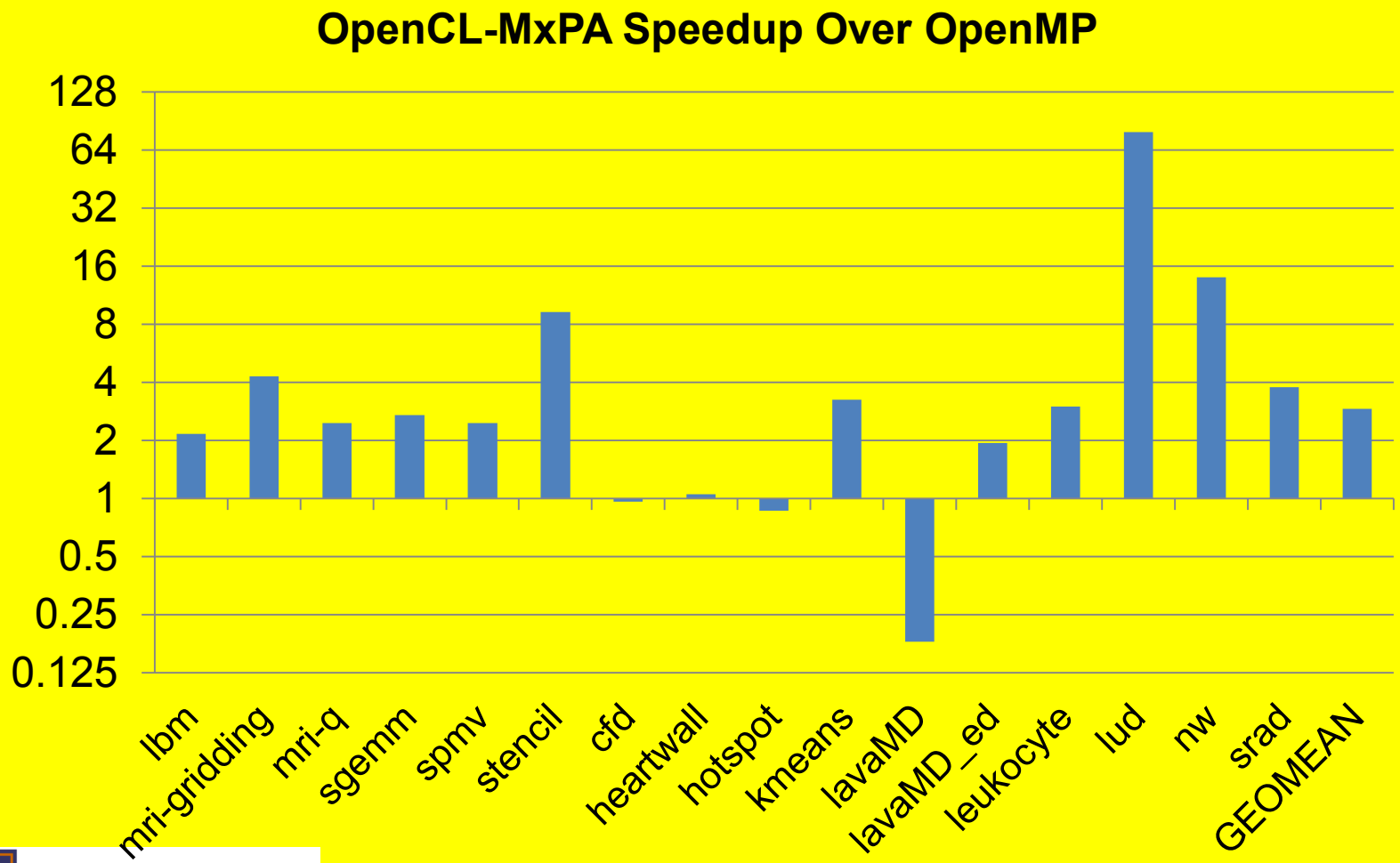
Data transfers in big-data analytics are even worse.



Actual Data Transfer Behavior



A single interface for both CPUs and GPUs is possible.



Stratton, Ph.D. Thesis

OPTIMIZING ACROSS INTERFACES

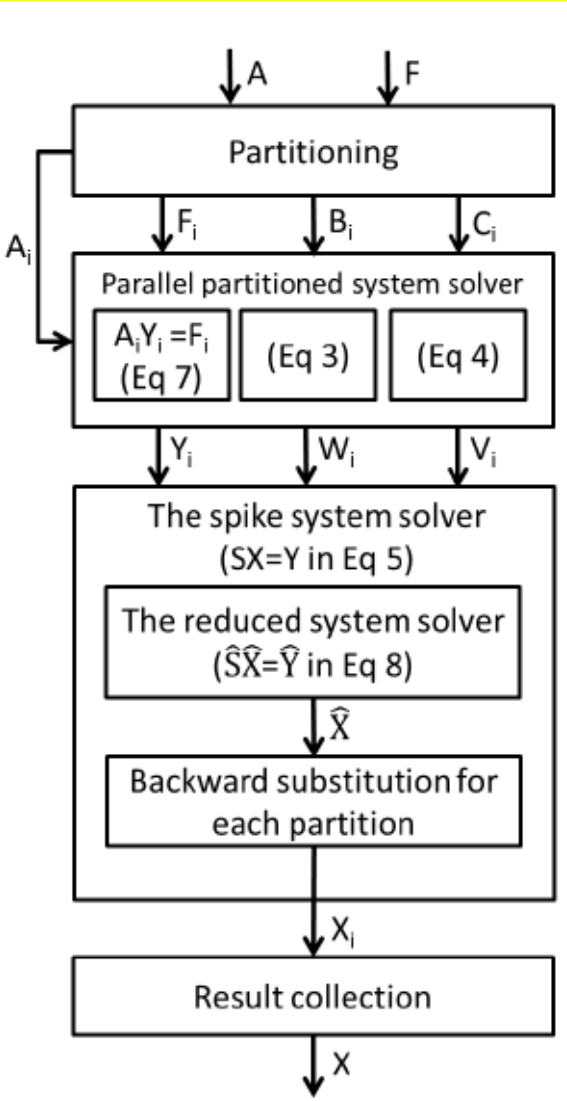
Numerically Stable Partitioning Algorithm

- SPIKE (Polizzi et al), diagonal pivoting for numerical stability

$$A=\begin{bmatrix} A_1 & B_1 & & \\ C_2 & A_2 & B_2 & \\ & C_3 & A_3 & B_3 \\ & & C_4 & A_4 \end{bmatrix}, \quad X=\begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix}, \quad F=\begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \end{bmatrix}.$$

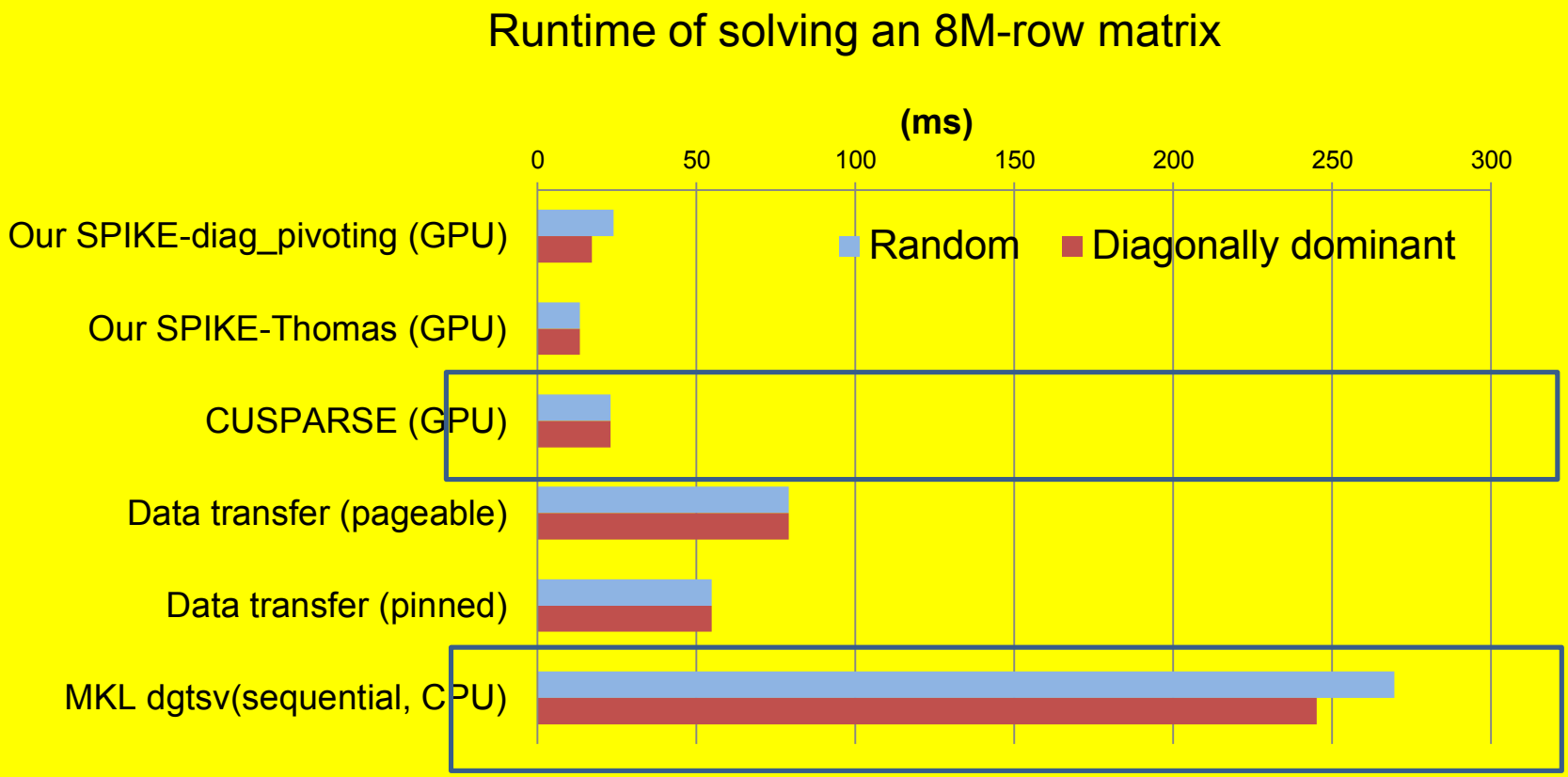
$$D=\begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & A_3 & \\ & & & A_4 \end{bmatrix}, \quad S=\begin{bmatrix} I & V_1 & & \\ W_2 & I & V_2 & \\ & W_3 & I & V_3 \\ & & W_4 & I \end{bmatrix}.$$

$$A_iV_i=\begin{bmatrix} 0 \\ \vdots \\ 0 \\ B_i \end{bmatrix} \quad (3) \qquad A_iW_i=\begin{bmatrix} C_i \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (4)$$



SX = Y (5) DY = F (6) AiYi = Fi (7)

GPU Performance Advantage

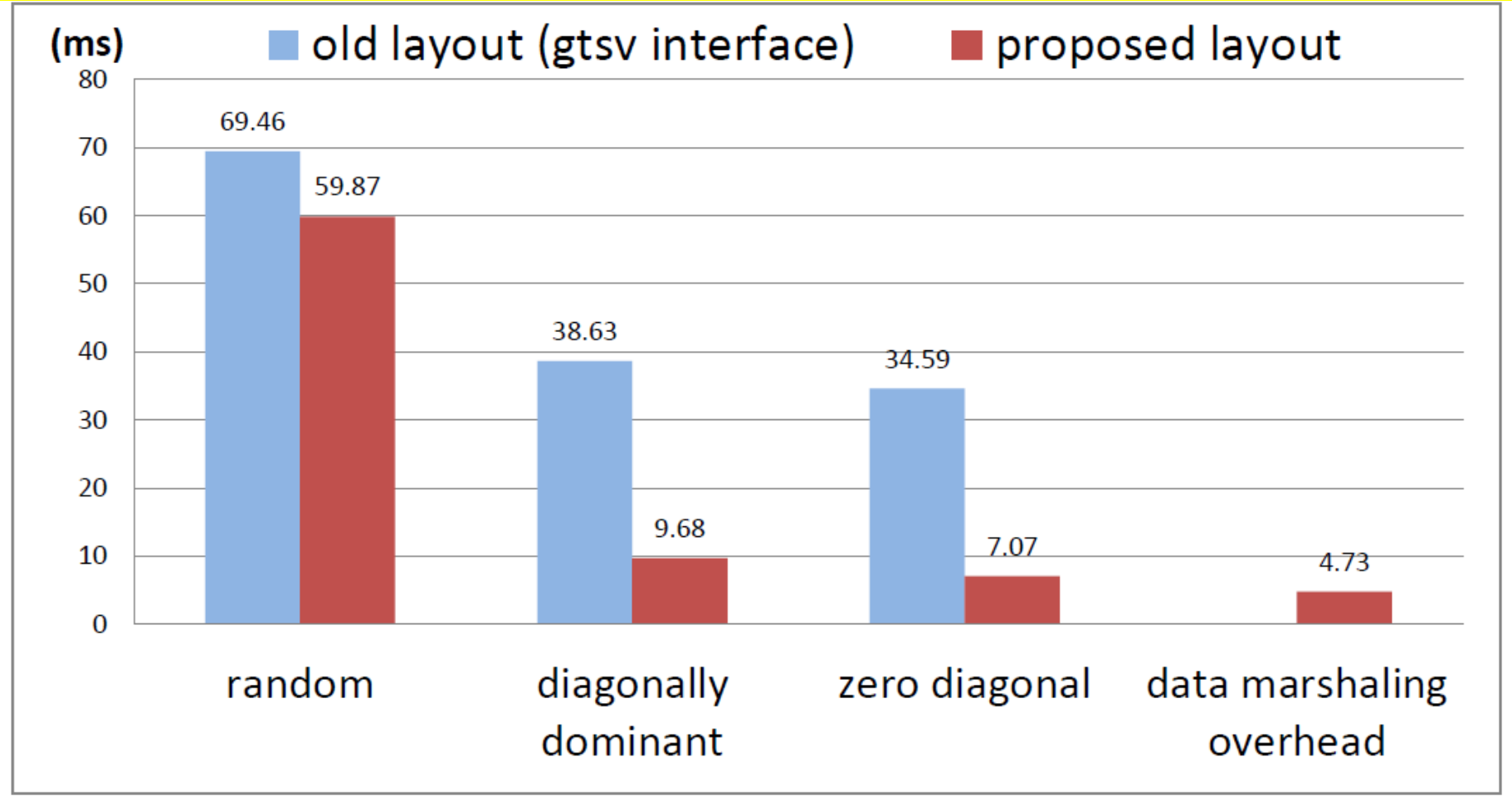


Numerical Error and Stability

Relative Backward Error

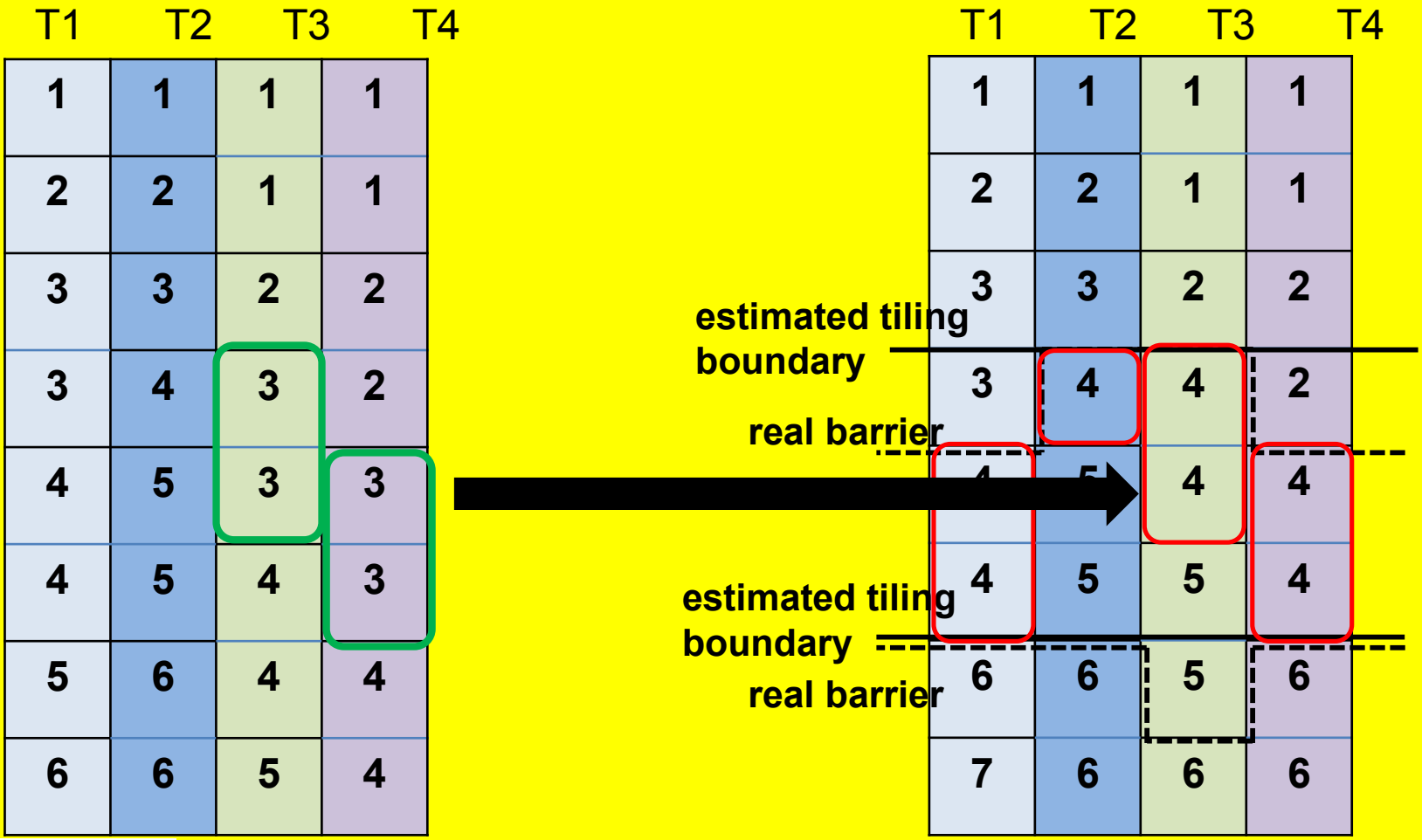
Matrix type	SPIKE-diag_pivoting	SPIKE-Thoma	CUSPARSE	MKL	ntel SPIKE	Matlab
1	1.82E-14	1.97E-14	7.14E-12	1.88E-14	1.39E-15	1.96E-14
2	1.27E-16	1.27E-16	1.69E-16	1.03E-16	1.02E-16	1.03E-16
3	1.55E-16	1.52E-16	2.57E-16	1.35E-16	1.29E-16	1.35E-16
4	1.37E-14	1.22E-14	1.39E-12	3.10E-15	1.69E-15	2.78E-15
5	1.07E-14	1.13E-14	1.82E-14	1.56E-14	4.62E-15	2.93E-14
6	1.05E-16	1.06E-16	1.57E-16	9.34E-17	9.51E-17	9.34E-17
7	2.42E-16	2.46E-16	5.13E-16	2.52E-16	2.55E-16	2.27E-16
8	2.14E-04	2.14E-04	1.50E+10	3.76E-04	2.32E-16	2.14E-04
9	2.32E-05	3.90E-04	1.93E+08	3.15E-05	9.07E-16	1.19E-05
10	4.27E-05	4.83E-05	2.74E+05	3.21E-05	4.72E-16	3.21E-05
11	7.52E-04	6.59E-02	4.54E+11	2.99E-04	2.20E-15	2.28E-04
12	5.58E-05	7.95E-05	5.55E-04	2.24E-05	5.52E-05	2.24E-05
13	5.51E-01	5.45E-01	1.12E+16	3.34E-01	3.92E-15	3.08E-01
14	2.86E+49	4.49E+49	2.92E+51	1.77E+48	3.86E+54	1.77E+48
15	2.09E+60	Nan	Nan	1.47E+59	Fail	3.69E+58
16	Inf	Nan	Nan	Inf	Fail	4.68E+171

Fast Transposition on GPU



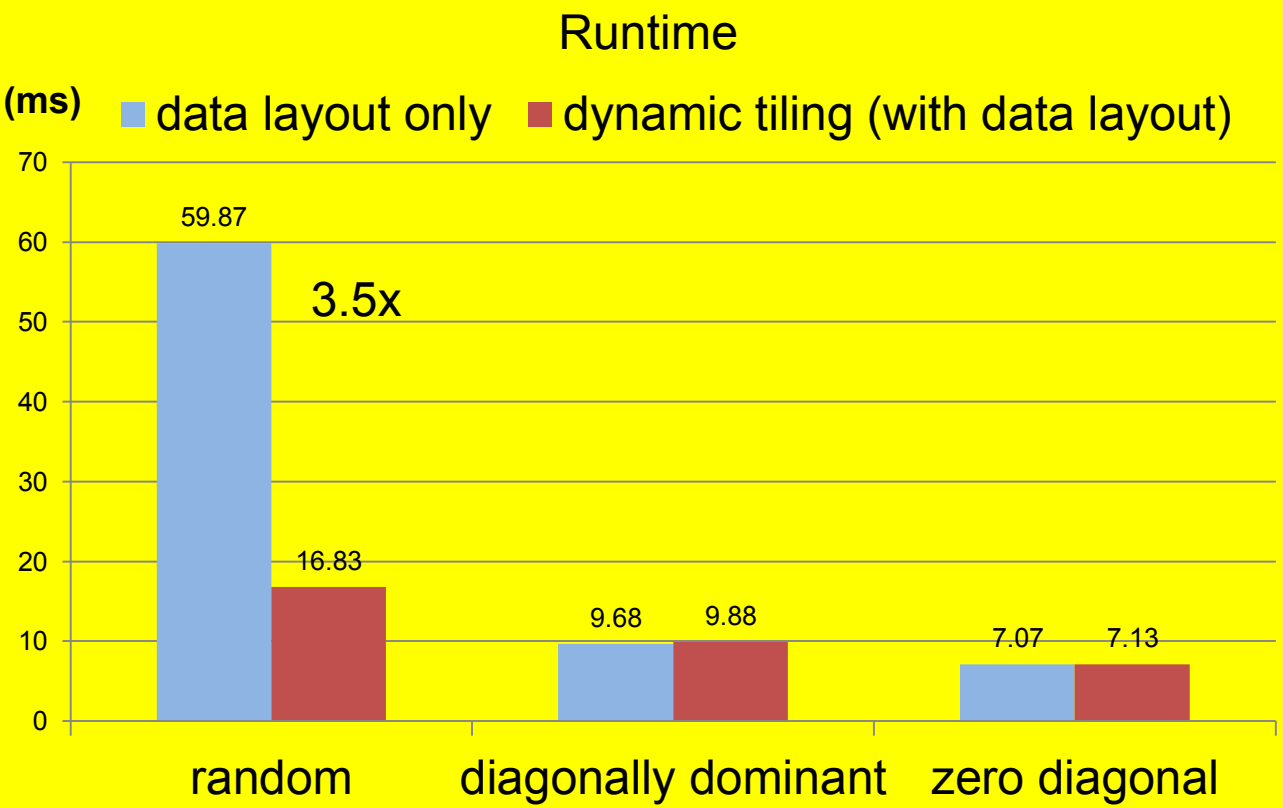
Chang, et al, SC 2012

Dynamic Tiling



Chang, et al, SC 2012

Dynamic Tiling



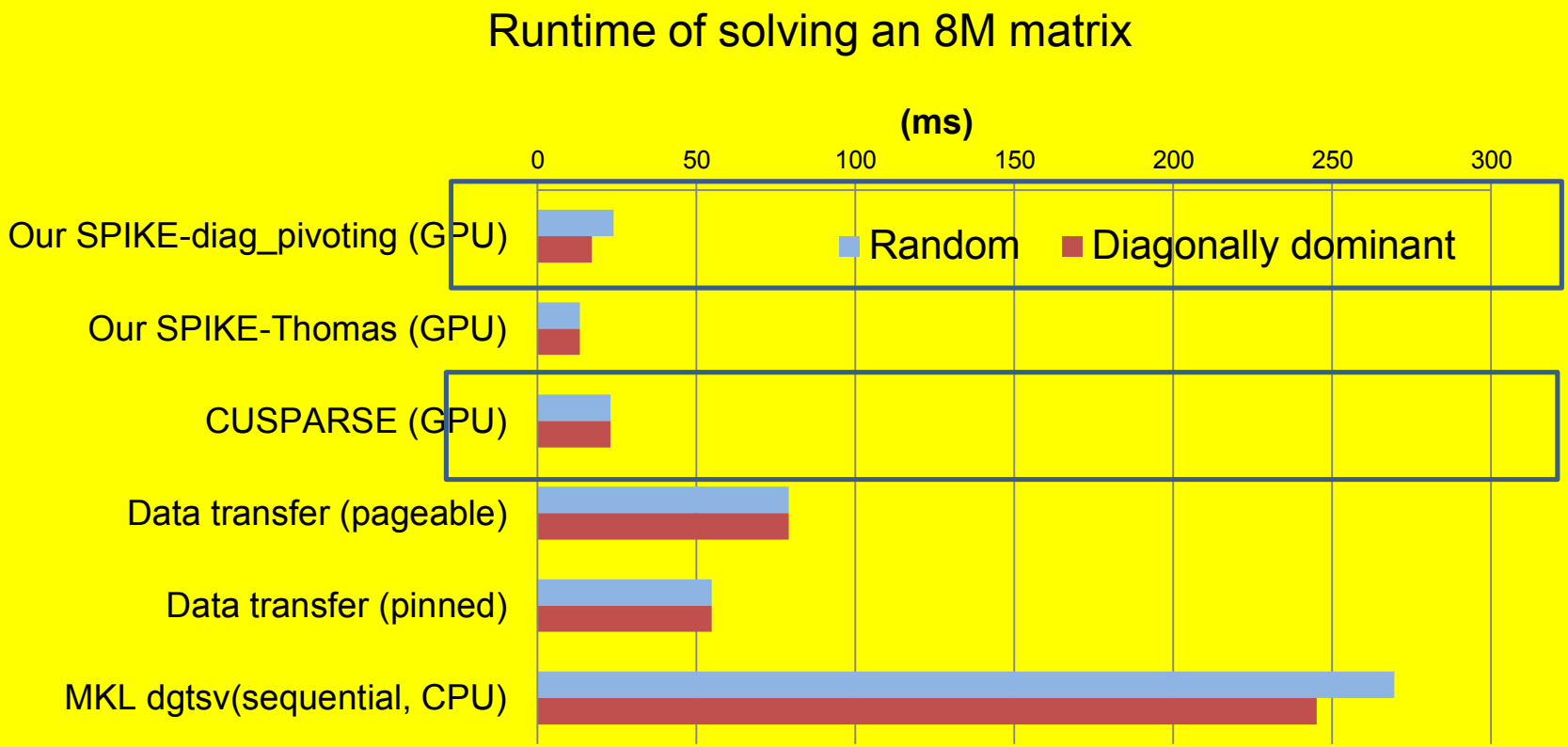
Numerical Error and Stability

Relative Backward Error

Matrix type	SPIKE-diag_pivoting	SPIKE-Thomas	CUSPARSE	MKL	Intel SPIKE	Matlab
1	1.82E-14	1.97E-14	7.14E-12	1.88E-14	1.39E-15	1.96E-14
2	1.27E-16	1.27E-16	1.69E-16	1.03E-16	1.02E-16	1.03E-16
3	1.55E-16	1.52E-16	2.57E-16	1.35E-16	1.29E-16	1.35E-16
4	1.37E-14	1.22E-14	1.39E-12	3.10E-15	1.69E-15	2.78E-15
5	1.07E-14	1.13E-14	1.82E-14	1.56E-14	4.62E-15	2.93E-14
6	1.05E-16	1.06E-16	1.57E-16	9.34E-17	9.51E-17	9.34E-17
7	2.42E-16	2.46E-16	5.13E-16	2.52E-16	2.55E-16	2.27E-16
8	2.14E-04	2.14E-04	1.50E+10	3.76E-04	2.32E-16	2.14E-04
9	2.32E-05	3.90E-04	1.93E+08	3.15E-05	9.07E-16	1.19E-05
10	4.27E-05	4.83E-05	2.74E+05	3.21E-05	4.72E-16	3.21E-05
11	7.52E-04	6.59E-02	4.54E+11	2.99E-04	2.20E-15	2.28E-04
12	5.58E-05	7.95E-05	5.55E-04	2.24E-05	5.52E-05	2.24E-05
13	5.51E-01	5.45E-01	1.12E+16	3.34E-01	3.92E-15	3.08E-01
14	2.86E+49	4.49E+49	2.92E+51	1.77E+48	3.86E+54	1.77E+48
15	2.09E+60	Nan	Nan	1.47E+59	Fail	3.69E+58
16	Inf	Nan	Nan	Inf	Fail	4.68E+171

Chang, et al, SC 2012

GPU Performance Advantage



Chang, et al, SC 2012

Computer Architecture as an Academic Discipline

- Few computer architects are among the top 20 most cited computer scientists
- Modest citation counts for ISCA papers
- Few top institutions consider Computer Architecture as a high-priority hiring area
- Few venture capitalists consider computer architecture as a major source of next big startups

Some Potential Actions

- Take charge of the major application programming interfaces
- Teach programmers what they need to know about computer architecture
- Encourage computer architecture students to understand more about applications and numerical methods
- Rethink the scope of computer architecture conferences

Acknowledgements

- D. August (Princeton), S. Baghsorkhi (Illinois), N. Bell (NVIDIA), D. Callahan (Microsoft), J. Cohen (NVIDIA), B. Dally (Stanford), J. Demmel (Berkeley), P. Dubey (Intel), M. Frank (Intel), M. Garland (NVIDIA), Isaac Gelado (BSC), M. Gschwind (IBM), R. Hank (Google), J. Hennessy (Stanford), P. Hanrahan (Stanford), M. Houston (AMD), T. Huang (Illinois), D. Kaeli (NEU), K. Keutzer (Berkeley), I. Gelado (UPC), B. Gropp (Illinois), D. Kirk (NVIDIA), D. Kuck (Intel), S. Mahlke (Michigan), T. Mattson (Intel), N. Navarro (UPC), J. Owens (Davis), D. Padua (Illinois), S. Patel (Illinois), Y. Patt (Texas), D. Patterson (Berkeley), C. Rodrigues (Illinois), S. Ryoo (ZeroSoft), K. Schulten (Illinois), B. Smith (Microsoft), M. Snir (Illinois), I. Sung (Illinois), P. Stenstrom (Chalmers), J. Stone (Illinois), S. Stone (Harvard) J. Stratton (Illinois), H. Takizawa (Tohoku), M. Valero (UPC)

- And many others!

There is always hope.

– Aragorn in the eve of the Battle of Pelennor Minas Tirith

THANK YOU!