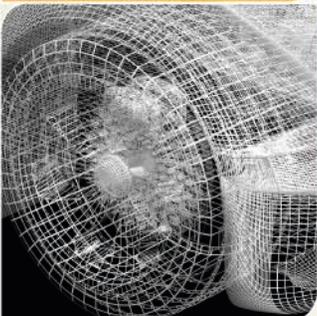


Accelerators, Cell Broadband Engine, Graphics Processors, and FPGAs

David A. Bader



Georgia Tech  College of Computing
Computational Science and Engineering

Georgia Tech's new School of Computational Science and Engineering



- Creation of the Computational Science and Engineering dept. and academic programs
- Klaus Advanced Computing Building (opened 2007)
- Prominent faculty hires in HPC
- GA Tech top ranked academic institution in a recent Top500 List of Supercomputing Sites
 - Recently deployed ~25 Teraflops
- Sony-Toshiba-IBM Cell Center of Competence
- Center for Advanced Supercomputing Software (CASS-MT)
- NSF IUCRC for Multicore Computing Productivity Research (CMPR)
- NSF IUCRC for Experimental Computing Systems (CERCS)
- Oak Ridge joint faculty
- NSF PetaApps awards, several TeraGrid users

CRAY

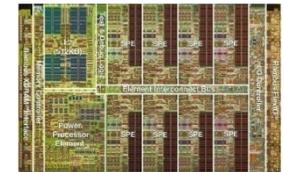
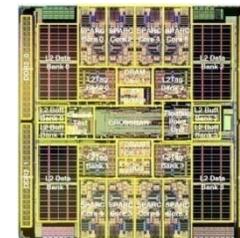
SONY



IBM



intel Leap ahead



nvidia

TOSHIBA



NATIONAL INSTITUTES OF HEALTH



David A. Bader



DARPA

OAK RIDGE National Laboratory



Pacific Northwest NATIONAL LABORATORY

Sandia National Laboratories

Georgia Tech



Microsoft Research



College of Computing

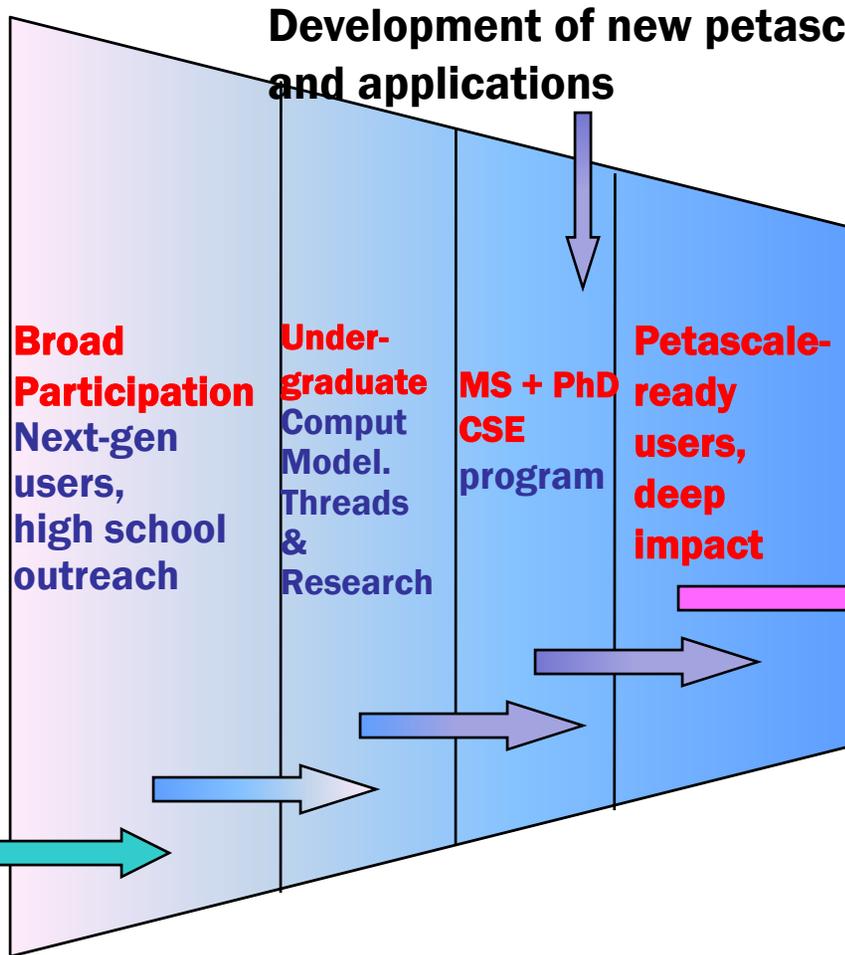
Georgia Tech's CSE creates a Petascale Pipeline to Accelerate New Science, Engineering, and Users



Petascale Foundation
Curricula, Training, Outreach



Broader participation and impact



Development of new petascale users and applications

Broad Participation
Next-gen users, high school outreach

Undergraduate
Comput Model. Threads & Research

MS + PhD
CSE program

Petascale-ready
users, deep impact

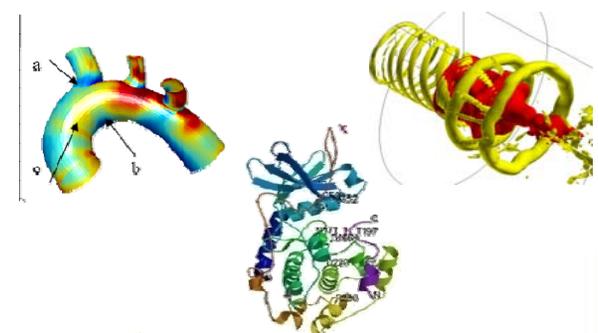
Petascale Systems

Breakthrough science and engineering at multiple levels

GA Tech grows the nation's pipeline



Dual Degrees Student from Morehouse College and GA Tech



David A. Bader

Broadening Participation in Petascale Science and Engineering

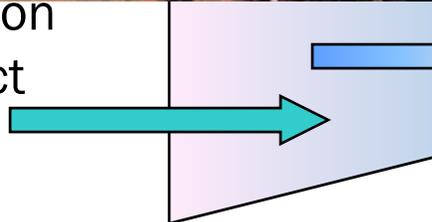


**Petascale
Foundation
Curricula,
Training,
Outreach**

**Broad
Participation**



B
participation
and impact



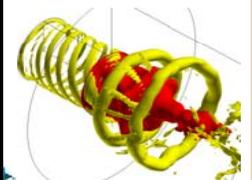
Curricula and outreach to increase the broad participation and impact

- **Education portal** for national access to all materials
- **Engagement with and enhancement materials for under-represented groups** (Intel Opportunity Scholars and SAIC Scholars promote undergraduate research experiences and help retain underrepresented minorities and women in computer science at GA Tech.
- **Summer Undergraduate Internships** with petascale faculty users such as the CRUISE (CSE Research for Undergraduates in Summer Experience) program, with support from NSF REU, the DoD HPC JOEM minority program, and industry.



Dual Degrees Student from Morehouse College and GA Tech

gh
d
y at
els



GA Tech grows the nation's pipeline

In aggregate, we provide a **comprehensive, pipelined environment**

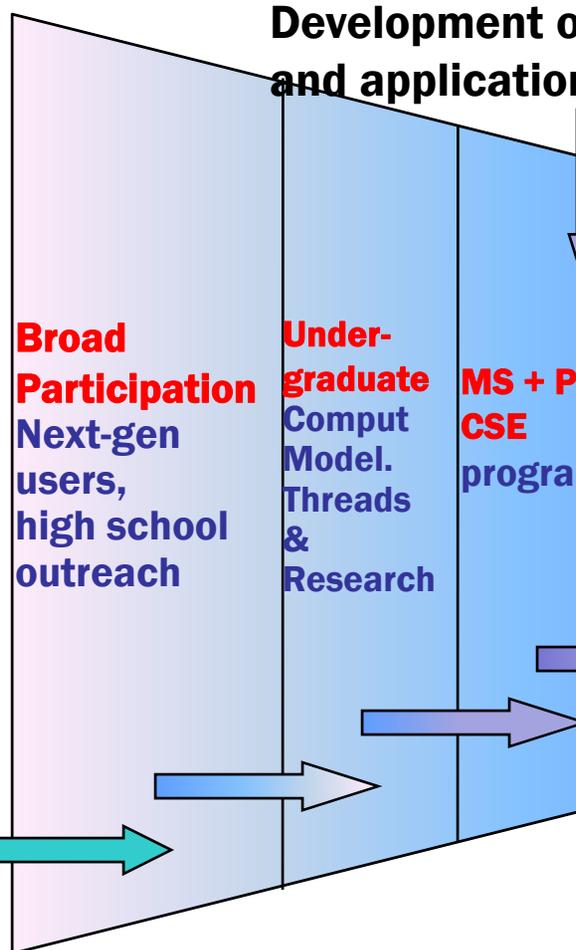
for petascale science and engineering



Petascale Foundation
Curricula, Training, Outreach



Broader participation and impact

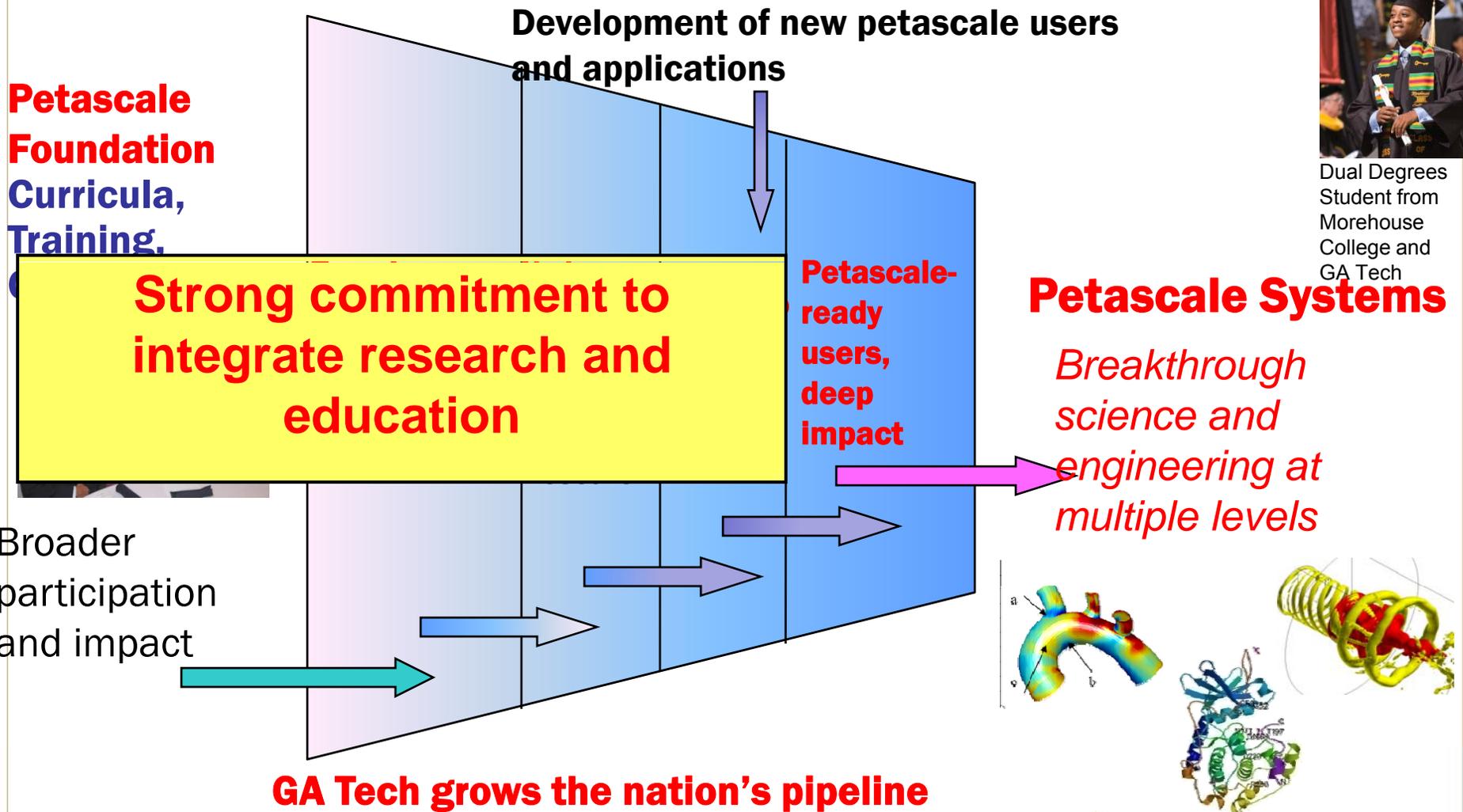


GA Tech grows the nation's

Curricula and education targeted to the next-gen of petascale users

- Undergraduate course materials and courses in “**computational modeling**” thread: parallel programming for multicore, large-scale parallelism
- GA Tech is establishing a rotating **CSE seminar series** and **undergraduate research program** with **Morehouse College** (A. Johnson), and **Spelman College** (A. Lawrence).
- **Leadership in Education**: new MS & PhD graduate curriculum in CSE with **HPC**, **large-scale data analysis**, **modeling & simulation**, **num. methods**, and **real-world algorithms**.
- **Curriculum sharing** for national access to all materials through: **tutorials** and **workshops** at TG and SC; **Computational Science Education Reference Desk (CSERD)**, an NSF-supported national resource library

Georgia Tech promotes petascale research and education



Dual Degrees Student from Morehouse College and GA Tech

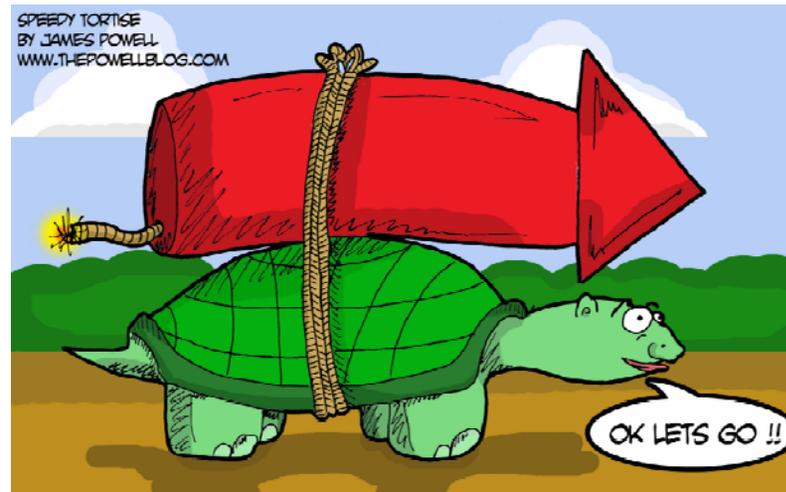


ac·cel·er·a·tor (noun)



- Pronunciation: \ik-'se-lə-'rā-tər, ak-\
Date: 1552

: one that accelerates: as **a**: a muscle or nerve that speeds the performance of an action **b**: a device (as a pedal) for controlling the speed of a motor vehicle engine **c**: a substance that speeds a chemical reaction **d**: an apparatus for imparting high velocities to charged particles (as electrons) **e**: an item of computer hardware that increases the speed at which a program or function operates <a graphics *accelerator*>





TiTech Tsubame cluster: 47 TF in Oct. 2006

- Tokyo Institute of Technology
- 655 nodes of Sun Fire x4600 with 16 Opteron cores and 1PB of high-density, high-perf storage
 - AMD Opteron 880/885 dual-core
- 360 ClearSpeed CSX600 Advance accelerators
- 2,304 ports of InfiniBand 10Gbps network
- Lustre, Grid
- (85 TF peak)



ClearSpeed

Military Supercomputer Sets Record

June 9, 2008, John Markoff

SAN FRANCISCO — An American military supercomputer, assembled from components originally designed for video game machines, has reached a long-sought-after computing milestone by processing more than 1.026 quadrillion calculations per second.

The new machine is more than twice as fast as the previous fastest supercomputer, the [I.B.M. BlueGene/L](#), which is based at [Lawrence Livermore National Laboratory](#) in California.

The new \$133 million supercomputer, called Roadrunner in a reference to the state bird of New Mexico, was devised and built by engineers and scientists at I.B.M. and [Los Alamos National Laboratory](#), based in Los Alamos, N.M. It will be used principally to solve classified military problems to ensure that the nation's stockpile of nuclear weapons will continue to work correctly as they age. The Roadrunner will simulate the behavior of the weapons in the first fraction of a second during an explosion.

Before it is placed in a classified environment, it will also be used to explore scientific problems like [climate change](#). The greater speed of the Roadrunner will make it possible for scientists to test global climate models with higher accuracy.

To put the performance of the machine in perspective, Thomas P. D'Agostino, the administrator of the National Nuclear Security Administration, said that if all six billion people on earth used hand calculators and performed calculations 24 hours a day and seven days a week, it would take them 46 years to do what the Roadrunner can in one day.

The machine is an unusual blend of chips used in consumer products and advanced parallel computing technologies. The lessons that computer scientists learn by making it calculate even faster are seen as essential to the future of both personal and mobile consumer computing.

The high-performance computing goal, known as a petaflop — one thousand trillion calculations per second — has long been viewed as a crucial milestone by military, technical and scientific organizations in the United States, as well as a growing group including Japan, China and the [European Union](#). All view supercomputing technology as a symbol of national economic competitiveness.

By running programs that find a solution in hours or even less time — compared with as long as three months on older generations of computers — petaflop machines like Roadrunner have the potential to fundamentally alter science and engineering, supercomputer experts say. Researchers can ask questions and receive answers virtually interactively and can perform experiments that would previously have been impractical.

"This is equivalent to the four-minute mile of supercomputing," said Jack Dongarra, a computer scientist at the [University of Tennessee](#) who for several decades has tracked the performance of the fastest computers.

Each new supercomputing generation has brought scientists a step closer to faithfully simulating physical reality. It has also produced software and hardware technologies that have rapidly spilled out into the rest of the computer industry for consumer and business products.

Technology is flowing in the opposite direction as well. Consumer-oriented computing began dominating research and development spending on technology shortly after the cold war ended in the late 1980s, and that trend is evident in the design of the world's fastest computers.

The Roadrunner is based on a radical design that includes 12,960 chips that are an improved version of an I.B.M. Cell microprocessor, a parallel processing chip originally created for [Sony's PlayStation 3](#) video-game machine. The Sony chips are used as accelerators, or turbochargers, for portions of calculations.

The Roadrunner also includes a smaller number of more conventional Opteron processors, made by [Advanced Micro Devices](#), which are already widely used in corporate servers.

"Roadrunner tells us about what will happen in the next decade," said Horst Simon, associate laboratory director for computer science at the Lawrence Berkeley National Laboratory. "Technology is coming from the consumer electronics market and the innovation is happening first in terms of cellphones and embedded electronics."

The innovations flowing from this generation of high-speed computers will most likely result from the way computer scientists manage the complexity of the system's hardware.

Roadrunner, which consumes roughly three megawatts of power, or about the power required by a large suburban shopping center, requires three separate programming tools because it has three types of processors. Programmers have to figure out how to keep all of the 116,640 processor cores in the machine occupied simultaneously in order for it to run effectively.

"We've proved some skeptics wrong," said Michael R. Anastasio, a physicist who is director of the Los Alamos National Laboratory. "This gives us a window into a whole new way of computing. We can look at phenomena we have never seen before."

Solving that programming problem is important because in just a few years personal computers will have microprocessor chips with dozens or even hundreds of processor cores. The industry is now hunting for new techniques for making use of the new computing power. Some experts, however, are skeptical that the most powerful supercomputers will provide useful examples.

"If Chevy wins the Daytona 500, they try to convince you the Chevy Malibu you're driving will benefit from this," said Steve Wallach, a supercomputer designer who is chief scientist of Convey Computer, a start-up firm based in Richardson, Tex.

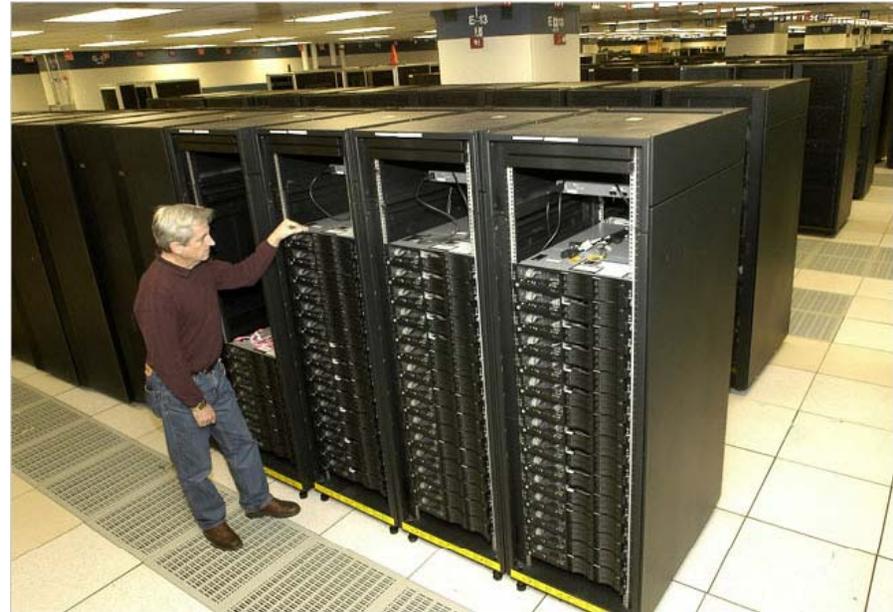
Those who work with weapons might not have much to offer the video gamers of the world, he suggested.

Many executives and scientists see Roadrunner as an example of the resurgence of the United States in supercomputing. Although American companies had dominated the field since its inception in the 1960s, in 2002 the Japanese Earth Simulator briefly claimed the title of the world's fastest by executing more than 35 trillion mathematical calculations per second. Two years later, a supercomputer created by I.B.M. reclaimed the speed record for the United States. The Japanese challenge, however, led Congress and the Bush administration to reinvest in high-performance computing.

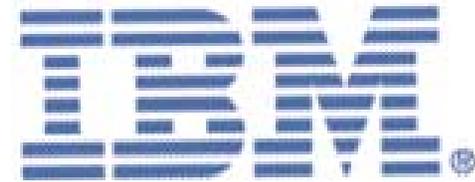
"It's a sign that we are maintaining our position," said Peter J. Ungaro, chief executive of Cray, a maker of supercomputers. He noted, however, that "the real competitiveness is based on the discoveries that are based on the machines."

Having surpassed the petaflop barrier, I.B.M. is already looking toward the next generation of supercomputing. "You do these record-setting things because you know that in the end we will push on to the next generation and the one who is there first will be the leader," said Nicholas M. Donofrio, an I.B.M. executive vice president.

By breaking the petaflop barrier sooner than had been generally expected, the United States' supercomputer industry has been able to sustain a pace of continuous performance increases, improving a thousandfold in processing power in 11 years. The next thousandfold goal is the exaflop, which is a quintillion calculations per second, followed by the zettaflop, the yottaflop and the xeraflop.



The Roadrunner supercomputer costs \$133 million and will be used to study nuclear weapons.



The New York Times

David A. Bader

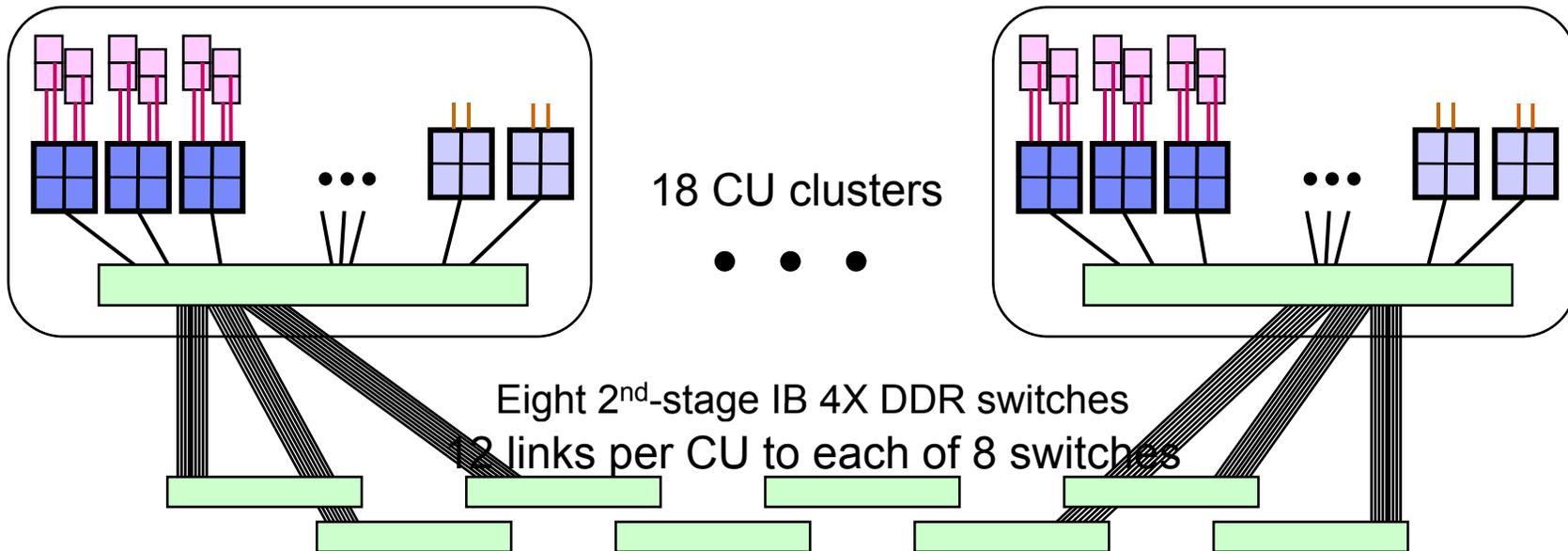
Roadrunner is a petascale system in 2008

Full Roadrunner Specifications:

6,912 dual-core Operons
 49.8 TF DP peak Operon
 27.6 TB Operon memory

12,960 Cell eDP chips
 1.33 PF DP peak Cell eDP
 2.65 PF SP peak Cell eDP
 51.8 TB Cell memory
 277 TB/s Cell memory BW

3,456 nodes on 2-stage IB 4X DDR
 13.8 TB/s aggregate BW (bi-dir) (1st stage)
 6.9 TB/s aggregate BW (bi-dir) (2nd stage)
 3.5 TB/s bi-section BW (bi-dir) (2nd stage)
 432 10 GigE I/O links on 216 I/O nodes
 432 GB/s aggregate I/O BW (uni-dir)
 (IB limited)





CPU vs. Accelerator

- How can I tell what is the CPU and what is the accelerator?
 - CPU is general-purpose, jack-of-all-trades
 - Accelerator is a specialized performance chip



- More and more the CPU is the controller and the accelerator is the dog



Single core / sequential computing





Cluster Computing





Use Multicore for Fault Tolerance?





Heterogeneous Manycore

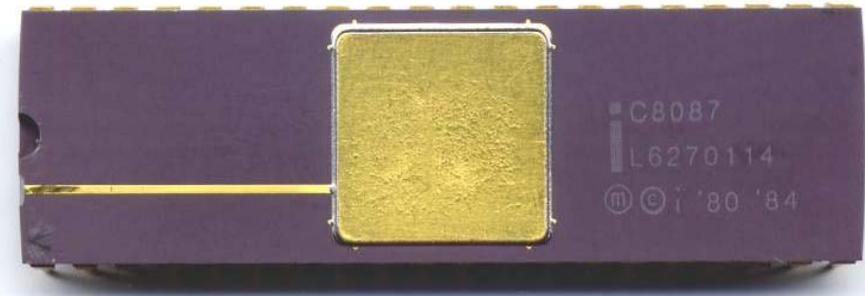


David A. Bader



Accelerators are not a new phenomenon

- **1980-1:** Intel introduced a co-processor
8088 CPU could invoke the 8087 accelerator
for fast fp operation
 - Intel C8087
 - 60 new instructions
 - 50 KF/s



... 80287 co-processor, 80387 co-processor

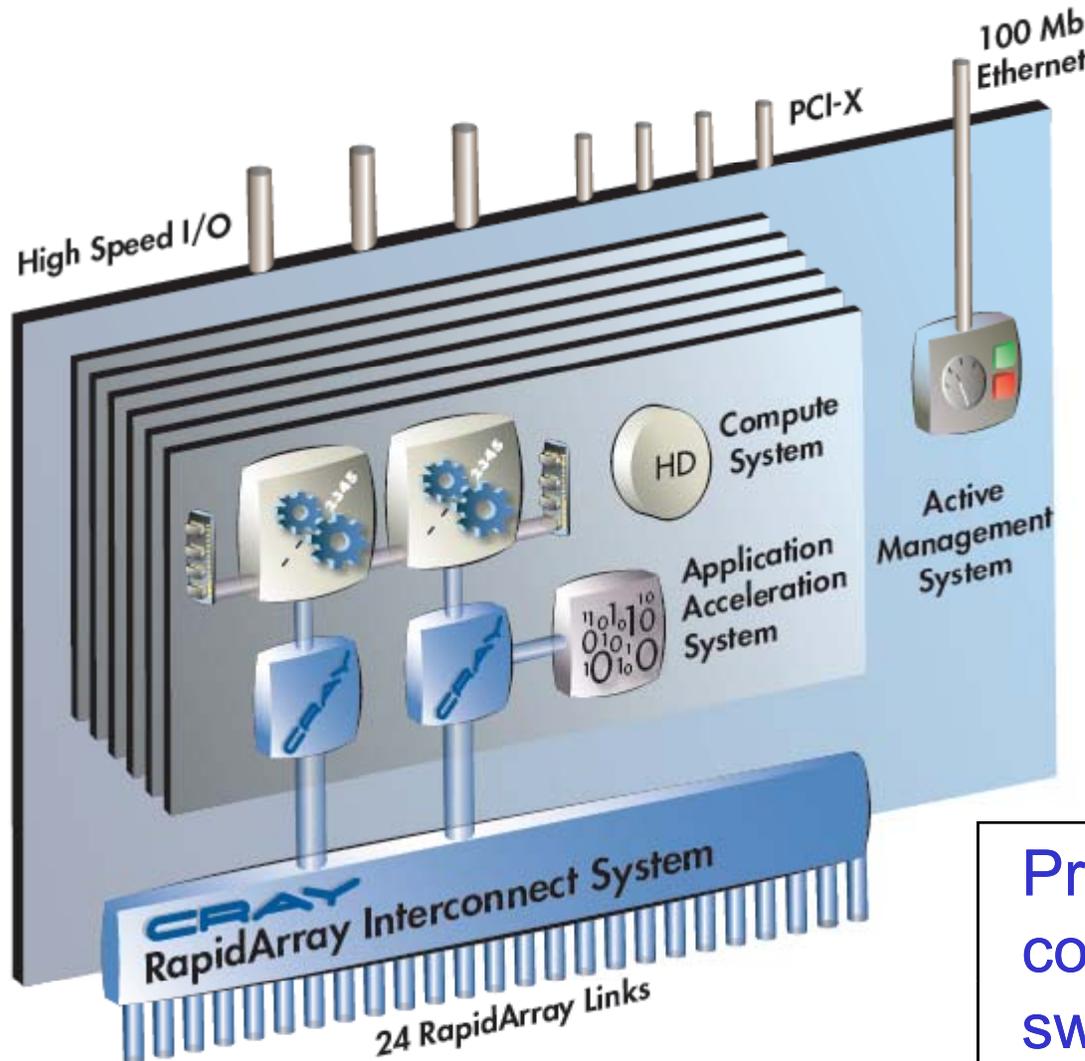
- **1989:** Intel 80486DX moves FP on-chip



Recent history

- Reconfigurable computing using FPGA accelerators
- For example:
 - Cray XD1
 - SRC Computers SRC-7

Cray XD1 System Architecture



Compute

- 12 AMD Opteron 32/64 bit, x86 processors
- High Performance Linux

RapidArray Interconnect

- 12 communications processors
- 1 Tb/s switch fabric

Active Management

- Dedicated processor

Application Acceleration

- 6 co-processors

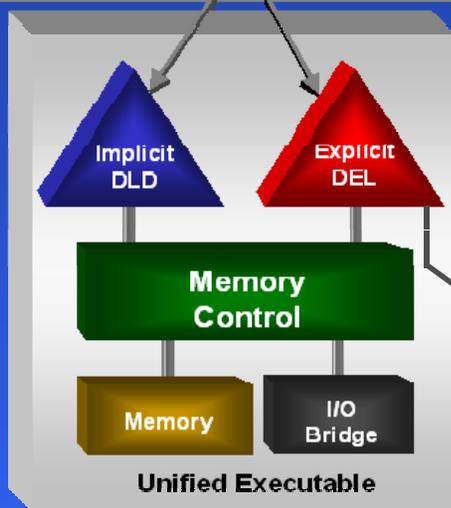
Processors directly connected via integrated switch fabric

SRC Architecture/MAP[®] Overview

Fortran → Carte™ Programming Environment ← C

• **Dense Logic Device (DLD)**

Implicit control
Higher clock rates
Typically fixed logic
μP, DSP, ASIC, etc.

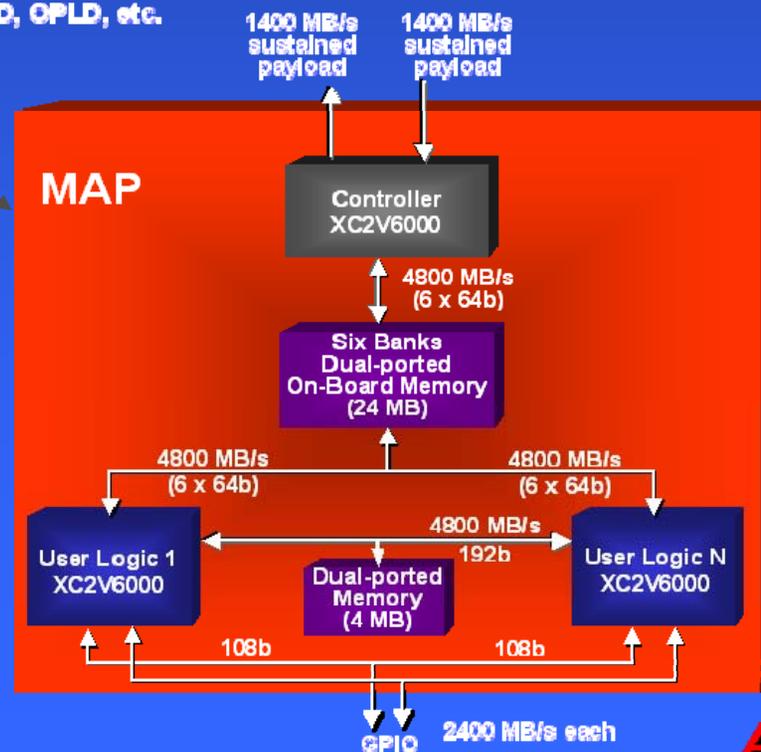


• **Direct Execution Logic (DEL)**

Explicit control
Lower clock rates
Typically reconfigurable
FPGA, CPLD, OPLD, etc.

Carte provides:

- C, Fortran parallel coding
- Mapping for DLD and/or DEL
- Manage IPC automatically
- A level of debug



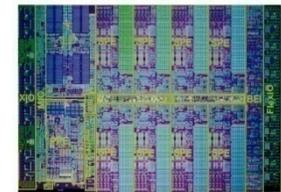
Limits to success with FPGA-

accelerated HPC systems

- Requires the user to partition the application between main CPU and accelerator
 - partitioning is NP-hard in the general case
- May lower productivity
 - Non-portable code
 - Non-deterministic results
 - Lack of IEEE 754 standard floating point
- Lack of elite programmers
 - Requires a computational scientist to develop HPC code and an electrical engineer to develop FPGA code

Today's potential accelerators and why they are so attractive

- FPGAs
 - Xilinx Virtex-5
- IBM Cell (Sony PlayStation 3)
 - 1+8 cores
 - PowerXCell 8i: 100GF/s dp (92W)
 - ALF/DaCS, OpenMP
- ClearSpeed: CSX700
 - 192 cores
 - 96 GF/s dp, 9W (typical)
 - SIMD array, Intel Math Kernel Library
- GPUs (graphics and entertainment)
 - nVidia GTX280: 78 GF/s dp, 236W
 - GPGPU, CUDA framework



ClearSpeed





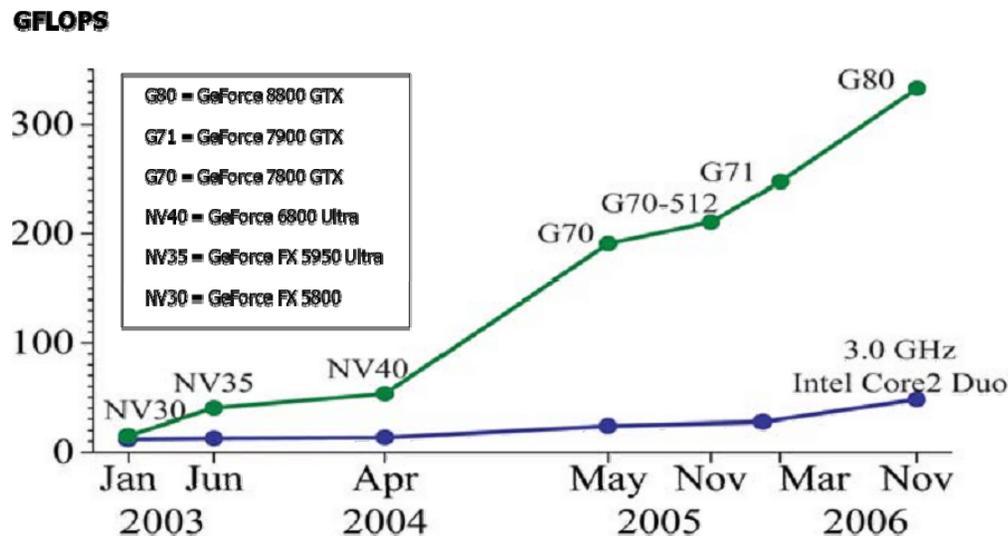
nVidia GeForce 8800



> 600M Transistors
 Excellent single-precision flops/\$
 $16 \times 8 = 128$ scalar fragment cores

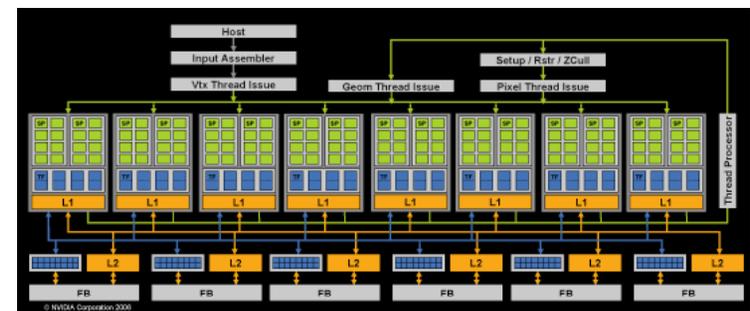
Streaming/SIMD processor
 Requires mapping application to shading/pixels

Some scientific computing maps well to GPU, for example:
 physics, FFT, sort



From "NVIDIA CUDA Programming Guide" version 0.8

David A. Bader





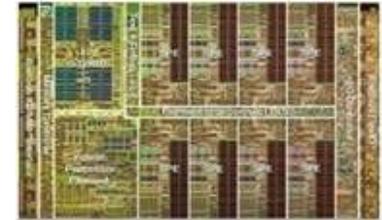
Problems with GPUs

- Must map problems to vertex shader, texture memory, rasterization
- Limited instruction set, missing key integer and bit operations
- CPU readback needed
- No indirect writes
- No native support for collective operations
 - reduce, gather, scatter, ...
- No support yet for GPU to GPU communication
- Requires higher-end *80 series for scientific computing
 - Binned by bit errors
- Programming
 - GPGPU, CUDA, OpenGL, DirectX

Sony-Toshiba-IBM Center of Competence for the Cell/B.E. at Georgia Tech



- ❖ **Mission:** grow the community of Cell Broadband Engine users and developers
- **Fall 2006:** Georgia Tech wins competition for hosting the STI Center
- First publicly-available IBM QS20 Cluster
- 200 attendees at 2007 STI Workshop
- Multicore curriculum and training
- **Demonstrated performance** on
 - Multimedia and gaming
 - Scientific computing
 - Medical applications
 - Financial services



David A. Bader, Director



TOSHIBA

IBM

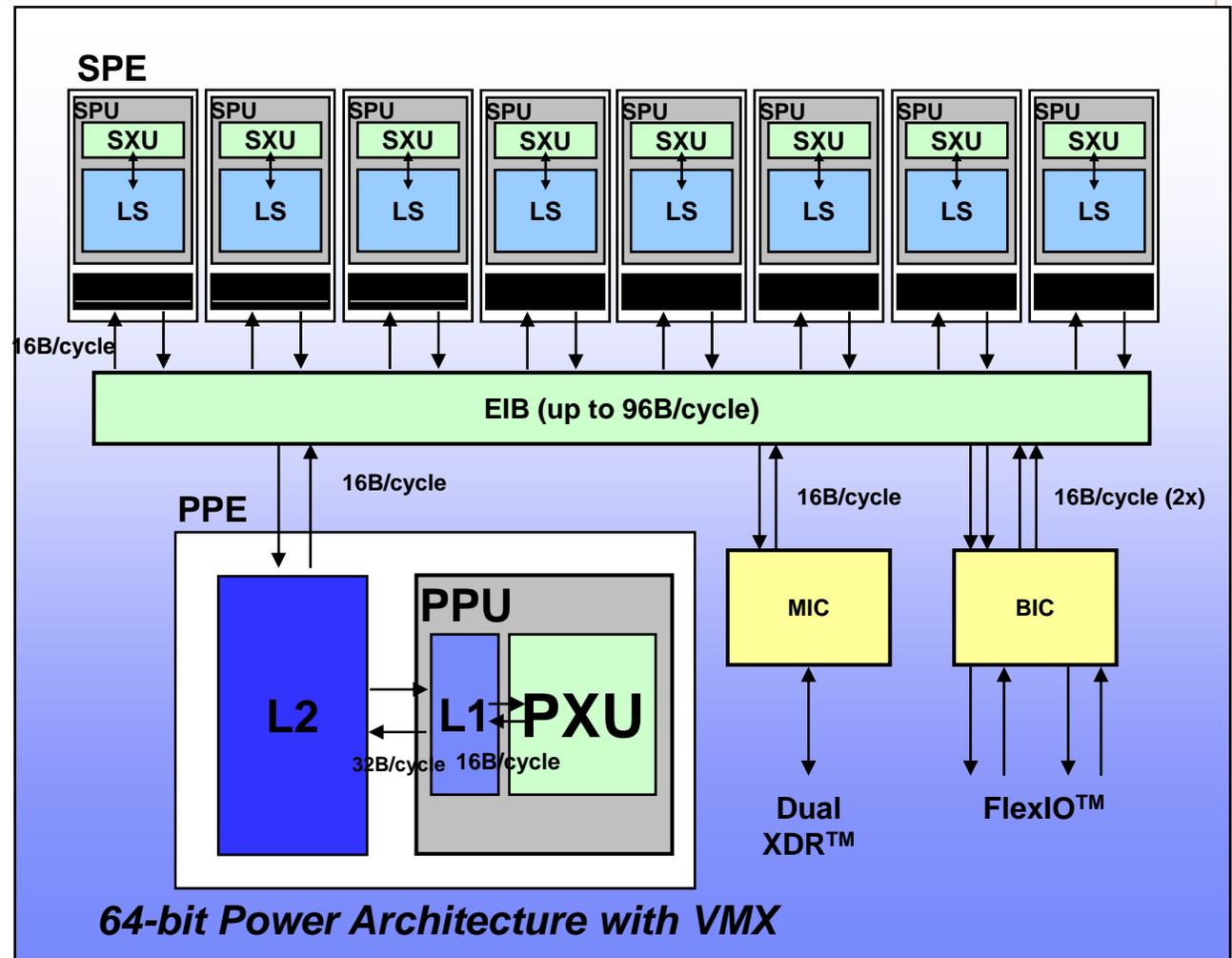
<http://sti.cc.gatech.edu>

David A. Bader



Cell BE Processor Features

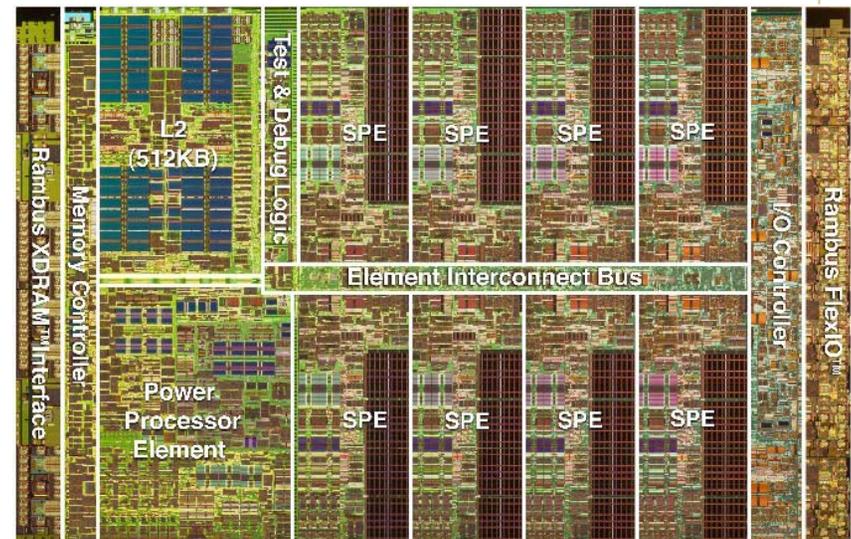
- Heterogeneous multi-core system architecture
 - Power Processor Element for control tasks
 - Synergistic Processor Elements for data-intensive processing
- Synergistic Processor Element (SPE) consists of
 - Synergistic Processor Unit (SPU)
 - Synergistic Memory Flow Control (SMF)
 - Data movement and synchronization
 - Interface to high-performance Element Interconnect Bus





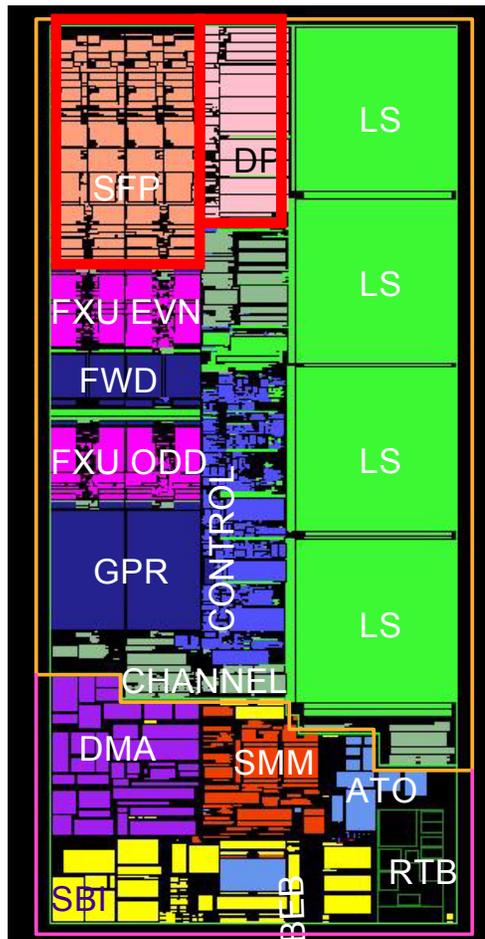
Cell BE Architecture

- Combines multiple high performance processors in one chip
 - 9 cores, 10 threads
 - A 64-bit Power Architecture™ core (PPE)
 - 8 Synergistic Processor Elements (SPEs) for data-intensive processing
- Current implementation—roughly 10 times the performance of Pentium for computational intensive tasks
 - Clock: 3.2 GHz (measured at >4GHz in lab)



	Cell	Pentium D
Peak I/O BW	75 GB/s	~6.4 GB/s
Peak SP Performance	~230 GFLOPS	~30 GFLOPS
Area	221 mm ²	206 mm ²
Total Transistors	234M	~230M

SPE Highlights

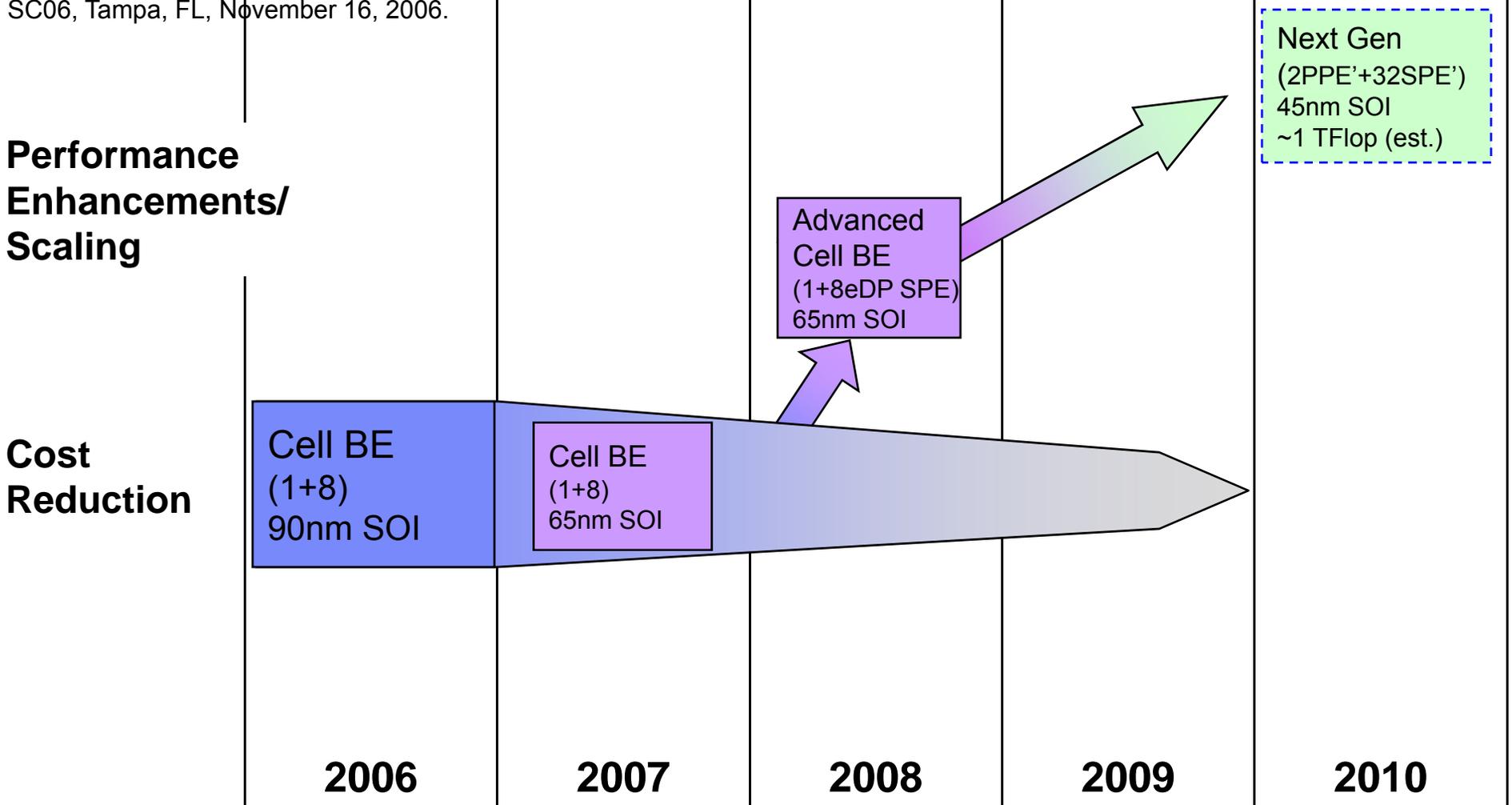


14.5mm² (90nm SOI)

- RISC like organization
 - 32 bit fixed instructions
 - Clean design – unified Register file
- User-mode architecture
 - No translation/protection within SPU
 - DMA is full Power Arch protect/x-late
- VMX-like SIMD dataflow
 - Broad set of operations (8 / 16 / 32 Byte)
 - Graphics SP-Float
 - IEEE DP-Float
- Unified register file
 - 128 entry x 128 bit
- 256KB Local Store
 - Combined I & D

Cell Broadband Engine Architecture™ (CBEA) Technology Competitive Roadmap

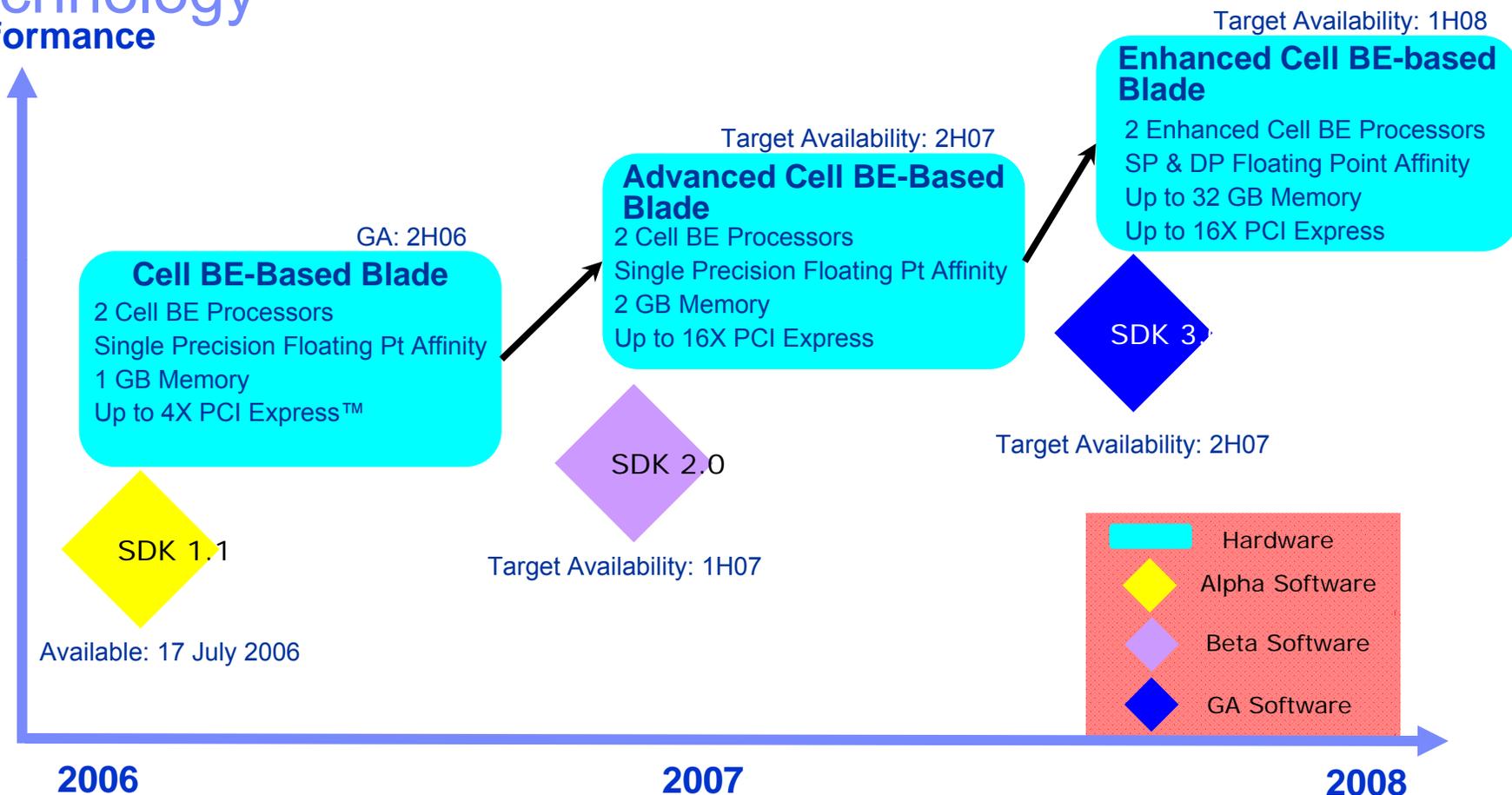
H. Peter Hofstee, *Real-time Supercomputing and Technology for Games and Entertainment*, Keynote Talk, SC06, Tampa, FL, November 16, 2006.



All future dates and specifications are estimations only; Subject to change without notice. Dashed outlines indicate concept designs.

Cell Broadband Engine™ Blade – The first in a line of planned offerings using Cell Broadband Engine technology

Performance



H. Peter Hofstee, *Real-time Supercomputing and Technology for Games and Entertainment*, Keynote Talk, SC06, Tampa, FL, November 16, 2006.

All future dates and specifications are estimations only; Subject to change without notice.



Conventional Wisdom and Reality

- 32-bit fp is 256GF/s peak (and 200 GF/s is achieved), but 64-bit fp is emulated → 25 GF/s
 - Today's PowerXCell 8i chips achieve 100 GF/s in native DP
- Programming for PPE and SPU, and hand-coding of DMA transfers, vectorization is difficult
 - Programming frameworks are available
 - RapidMind
 - Gadae Systems
 - Cells (UPC/BSC)
 - IBM XL C/C++ for Multicore Acceleration for Linux on {Power|x86} Systems V10.1, contains OpenMP Compiler for Cell/B.E.
- Gaming chips are power-hungry.
 - On the Green500 List, an IBM QS22 Cell system is #1 at 488.14 GF/W, while the best Intel solution is 265.81 GF/W
 - For DP, PowerXCell 8i is 3x more flops/watt than nVidia GTX280



Looking back: Success stories

- Good news! We know how to make acceleration work, for example:
 - IEEE 754 standard floating point
 - SSE/SSE2 for SIMD vectorization
 - Using GPU's for graphics (via OpenGL / DirectX)
- In each case, the accelerator is “hidden” from the typical end-user
- Heroic users can still work at the metal



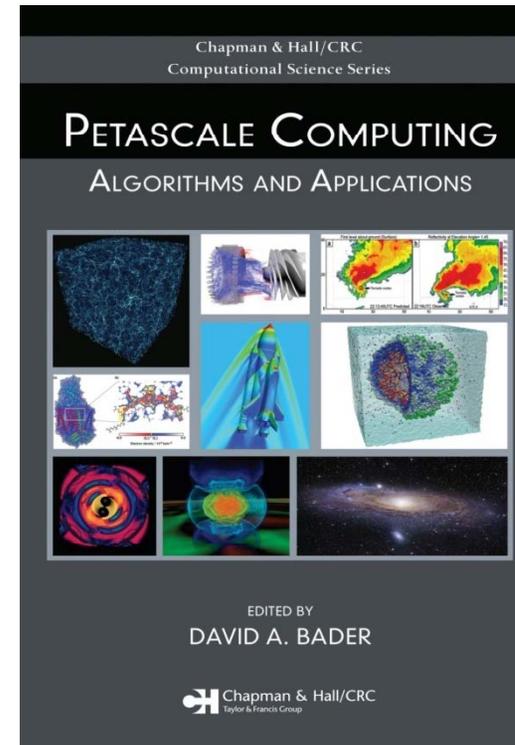
Looking forward: Future adoption

- Productivity is essential!
- Mass adoption of accelerators for computational science only will occur when the accelerator is seamlessly integrated, for example, through these strategies:
 - through the use of meta-compilers for standards (e.g. C+OpenMP) that handles CPU+accel. partitioning
 - through extension of the ISA, and
 - through portable libraries.

Petascale Computing: Algorithms and Applications (David A. Bader)



- Provides the first collection of articles on petascale algorithms and applications for computational science and engineering
- Covers a breadth of topics in petascale computing, including architectures, software, programming methodologies, tools, scalable algorithms, performance evaluation, and application development
- Discusses expected breakthroughs in the field for computational science and engineering
- Includes contributions from international researchers who are pioneers in designing applications for petascale computing systems



Chapman & Hall/CRC
Computational Science Series, © 2007



Acknowledgment of Support

