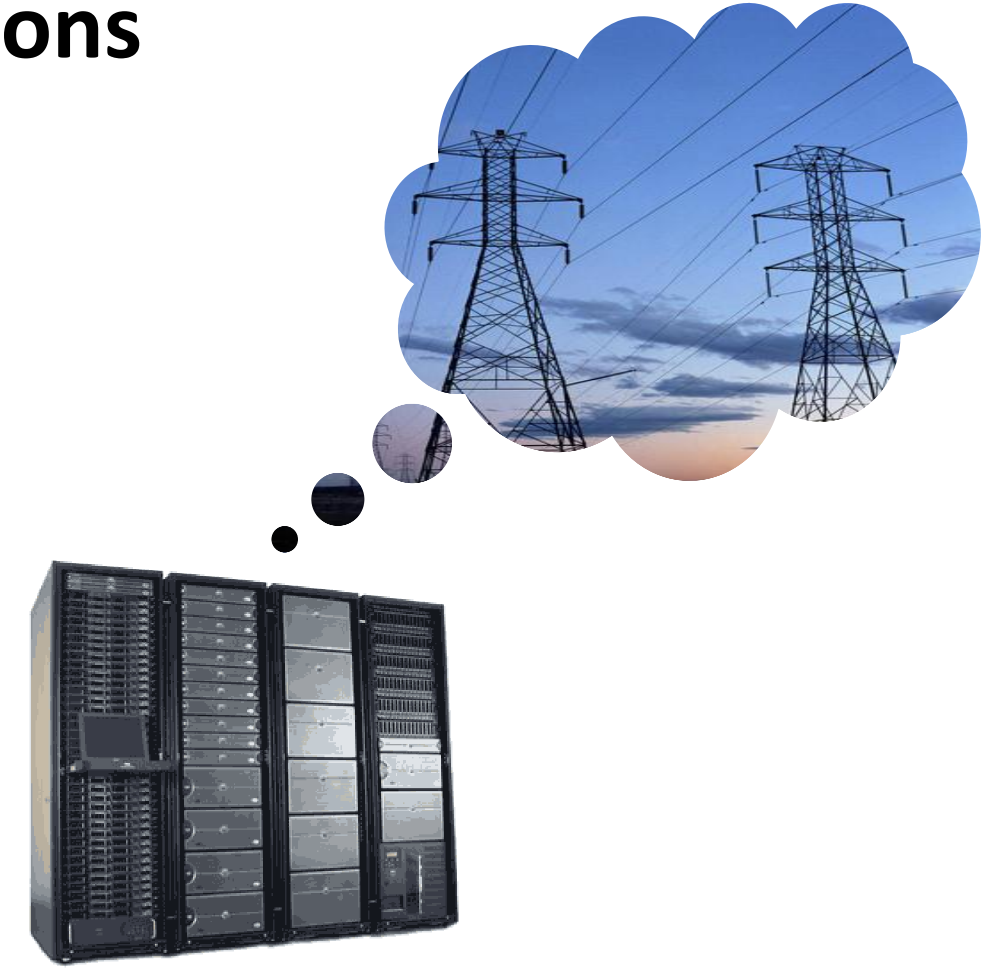
Trends in High-Performance Computing for Power Grid Applications

Franz Franchetti

ECE, Carnegie Mellon University
`www.spiral.net`

Co-Founder, SpiralGen
`www.spiralgen.com`



This work was supported by NSF-CPS award ECCS0931978

What Is High Performance Computing?

1 flop/s = one floating-point operation (addition or multiplication) per second
mega (M) = 10^6, giga (G) = 10^9, tera (T) = 10^{12}, peta (P) = 10^{15}, exa E) = 10^{18}

Computing systems in 2010

Cell phone	**Laptop**	**Workstation**	**HPC**	**#1 supercomputer**
1 CPUs	2 CPUs	8 CPUs	200 CPUs	224,162 CPUs
1 Gflop/s	20 Gflop/s	1 Tflop/s	20 Tflop/s	2.3 Pflop/s
$300	$1,200	$10,000	$700,000	$100,000,000
1 W power	30 W power	1 kW power	8 kW power	7 MW power

Economical HPC

Power grid scenario

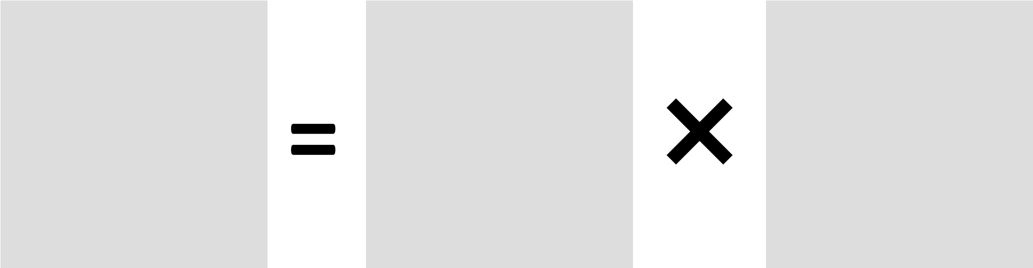
- Central servers (planning, contingency analysis)
- Autonomous controllers (smart grids)
- Operator workstations (decision support)

How Big are the Computational Problems?

1 flop/s = one floating-point operation (addition or multiplication) per second

mega (M) = 10^6, giga (G) = 10^9, tera (T) = 10^{12}, peta (P) = 10^{15}, exa E) = 10^{18}

Matrix-matrix multiplication…



```
for i=1:n
   for j=1:n
      for k=1:n
         C[i,j] +=
            A[i,k]*B[k,j]
```

..running on…

Cell phone

1 Gflop/s

1k × 1k

8MB, 2s

Laptop

20 Gflop/s

8k × 8k

0.5 GB, 5.5s

Workstation

1 Tflop/s

16k × 16k

2 GB, 8s

HPC

20 Tflop/s

64k × 64k

32 GB, 28s

#1 supercomputer

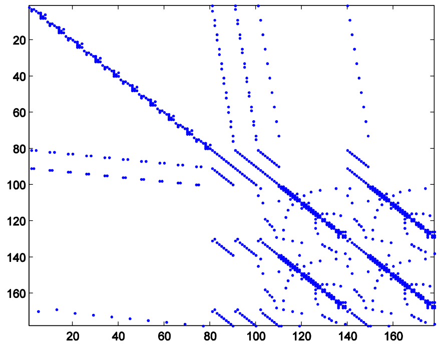
2.3 Pflop/s

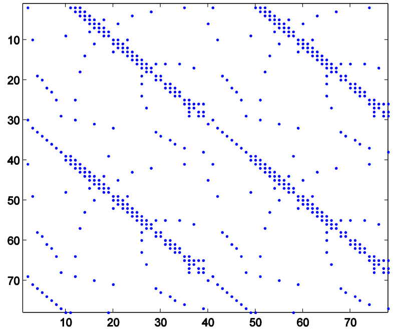
1M × 1M

8 TB, 1,000s

Matrix Sizes For Five Power Grids

	IEEE 14	New England 39	California	WSCC	Big Benchmark
			‘08 Heavy Summer	‘99 Heavy Summer	“Eastern Interconnect”
Bus number	14	39	4,138	11,855	24,000
Generator number	5	10	1070	2,149	8,000
Exciter number	5	10	910	1,630	8,000
Dynamic order	40	80	8,516	15,639	64,000
J_{sys}	78	178	18,932	43,647	104,000
J_{LF}	28 (23)	78 (67)	7,205	21,560	40,000
J_{R}	14	39	4,138	11,855	24,000
A	40	80	8,516	15,639	64,000





Matrices are *sparse*

Steady State Load Flow: Newton Raphson

Computation

"Solve non-linear load flow equations" with Newton Raphson iteration

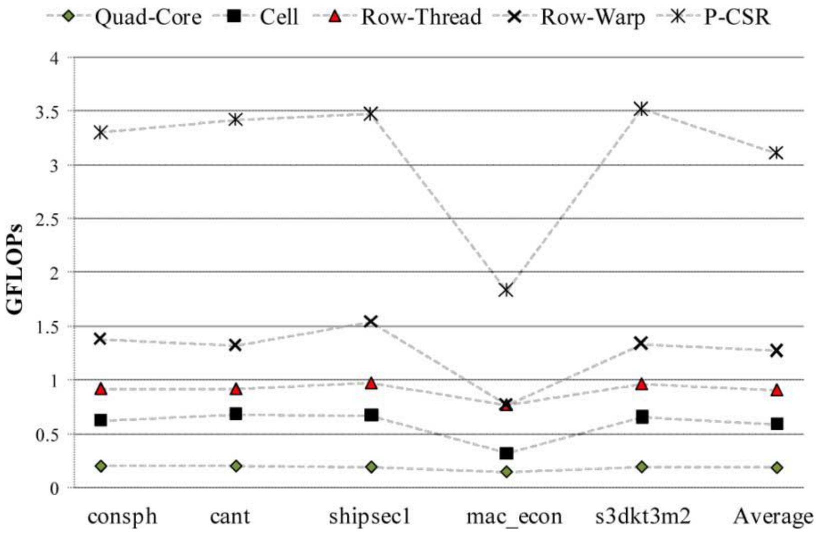
Kernel: iterative sparse linear solver (sparse LU)

Operations count: O($n^{1.4}$) per iteration

runtime proportional to memory bandwidth

Runtime

matrix size	time
20,000	0.3µs
100,000	3 µs
5,000,000	0.7s



Matrix Name	consph	cant	shipsec	mac_econ	s3dkt3m2
nnz	6010480	4007383	7813404	1273389	3843910
nnz/row	72.1	64.1	55.4	21.24	6.1
n=2	98.8	99.2	98	93.8	97.8
n=4	96.4	98	97.3	80	97.7
n=8	92.2	93.3	96.1	44	68.69

M. Mehri Dehnavi, D. Fernandez and D. Giannacopoulos: **Finite element sparse matrix vector multiplication on GPUs.** IEEE Transactions on Magnetics, vol. 46, no. 8, pp. 2982-2985, August 2010.

Image: Dell

50 GB/s

40 Gflop/s

Image: Nvidia

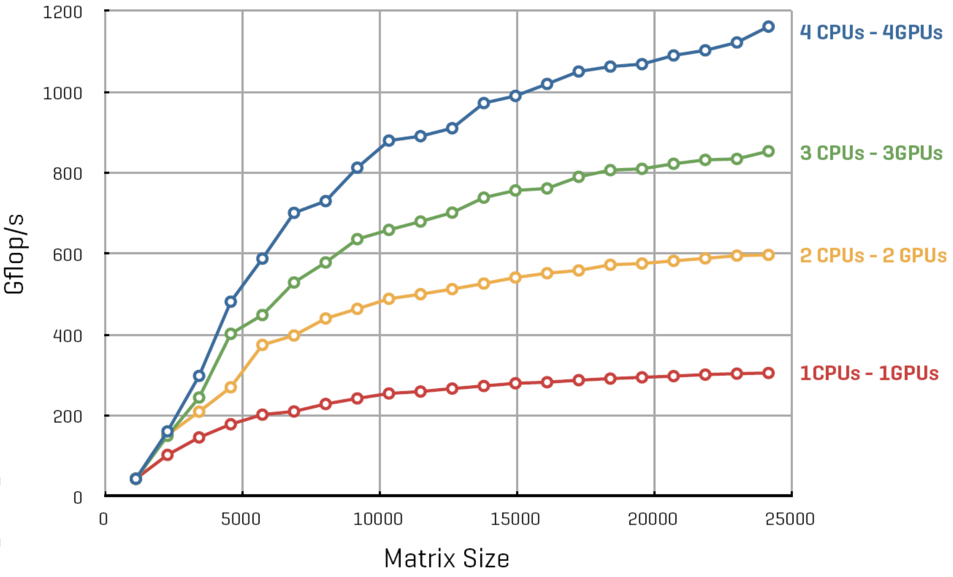
400 GB/s

80 Gflop/s

Voltage Stability: Dense Linear Algebra

Computation

Prove positive definiteness of J_R
“Run Cholesky factorization,
if successful, matrix is positive definite”
Operation count: n^3/3 flop for n x n matrix
LAPACK SPOTRF() computes $A=LL^T$



Runtime estimation

n	GB	Gflop	Gflop/s	t [s]
4,138	0.14	24	500	0.05
11,855	1.12	555	880	0.63
24,000	4.61	4,608	1,150	4.01

Server + GPU accelerator

4 Core i7 + 4 Tesla S1070
2.75 Tflop/s peak performance
$25,000 class

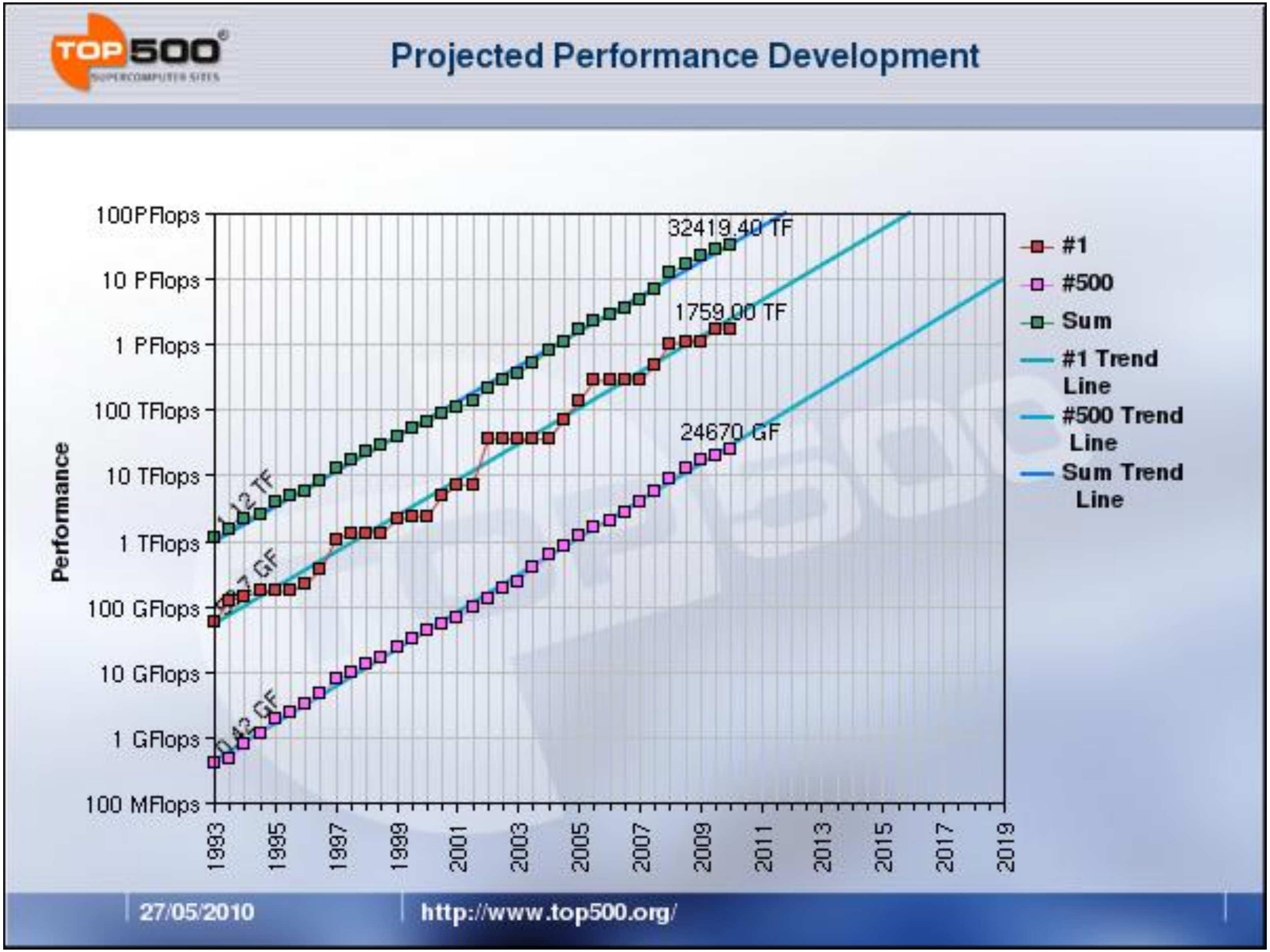
Image: Dell

+

Image: Nvidia

H. Ltaief, S. Tomov, R. Nath, P. Du, and Jack Dongarra
A Scalable High Performant Cholesky Factorization for Multicore with GPU Accelerators
LAPACK Working Note #223

The Evolution of Performance



How Do We Compare?

1 flop/s = one floating-point operation (addition or multiplication) per second

mega (M) = 10^6, giga (G) = 10^9, tera (T) = 10^{12}, peta (P) = 10^{15}, exa E) = 10^{18}

In 2010…

Cell phone
1 Gflop/s

Laptop
20 Gflop/s

Workstation
1 Tflop/s

HPC
20 Tflop/s

#1 supercomputer
2.3 Pflop/s

…would have been the #1 supercomputer back in…

Cray X-MP/48
941 Mflop/s
1984

NEC SX-3/44R
23.2 Gflop/s
1990

Intel ASCI Red
1.338 Tflop/s
1997

Earth Simulator
35.86 Tflop/s
2002

If History Predicted the Future…

…the performance of the #1 supercomputer of 2010…

#1 supercomputer
1 Pflop/s

…could be available as

HPC
1 Pflop/s
2018

How do we get here?

Workstation
1 Pflop/s
2023

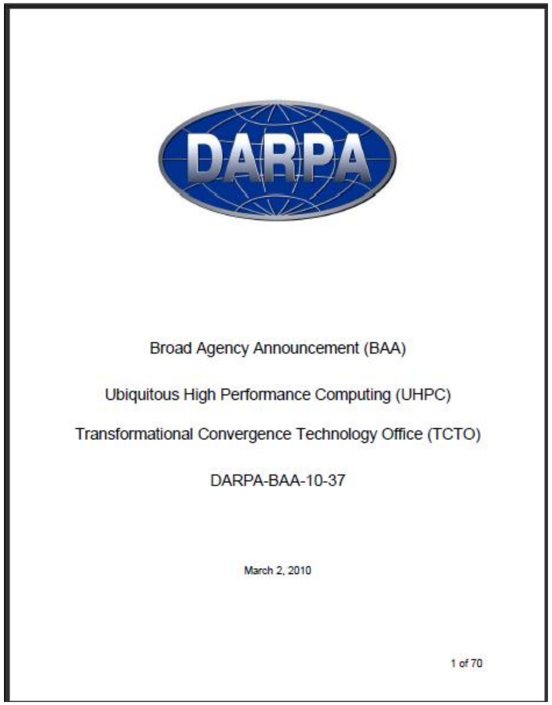
Laptop
1 Pflop/s
2030

Cell phone
1 Pflop/s
2036

HPC: ExtremeScale Computing

DARPA UHPC ExtremeScale system goals

- Time-frame: 2018
- 1 Pflop/s, air-cooled, single 19-inch cabinet
- Power budget: 57 kW, including cooling
- 50 Gflop/W for HPL benchmark



Broad Agency Announcement (BAA)

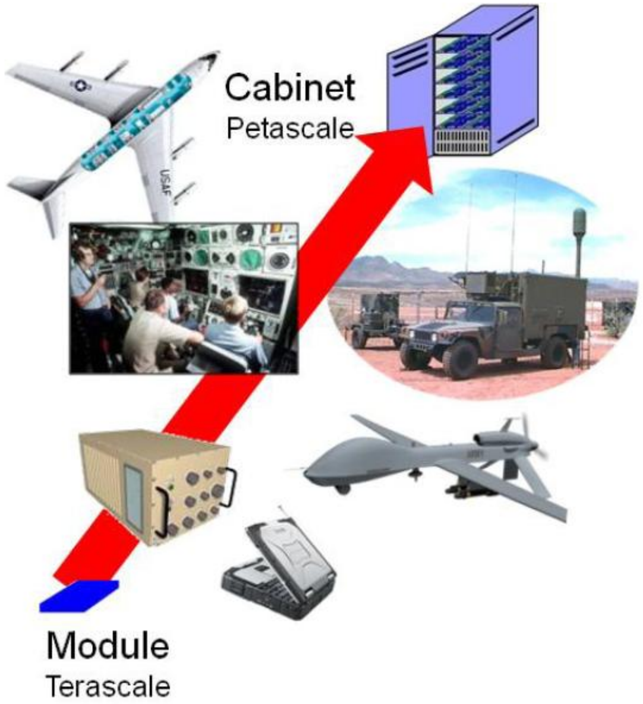
Ubiquitous High Performance Computing (UHPC)

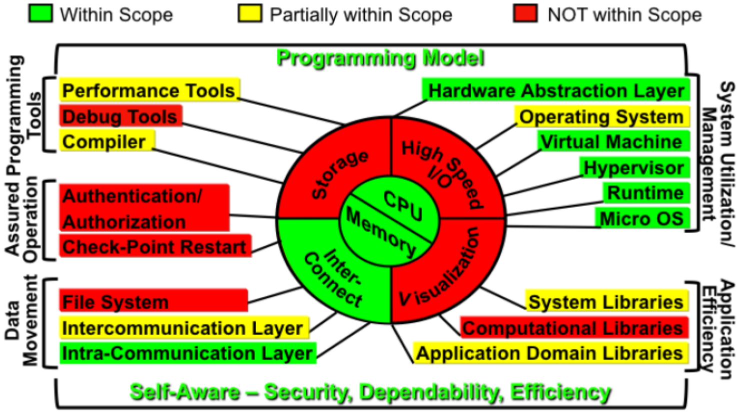
Transformational Convergence Technology Office (TCTO)

DARPA-BAA-10-37

March 2, 2010

1 of 70





🟩 Within Scope 🟨 Partially within Scope 🟥 NOT within Scope

Programming Model

Programming Tools: 🟨 Performance Tools; 🟥 Debug Tools; 🟨 Compiler

Assured Operation: 🟥 Authentication/ Authorization; 🟥 Check-Point Restart

Data Movement: 🟥 File System; 🟨 Intercommunication Layer; 🟩 Intra-Communication Layer

System Utilization/ Management: 🟩 Hardware Abstraction Layer; 🟨 Operating System; 🟩 Virtual Machine; 🟩 Hypervisor; 🟩 Runtime; 🟩 Micro OS

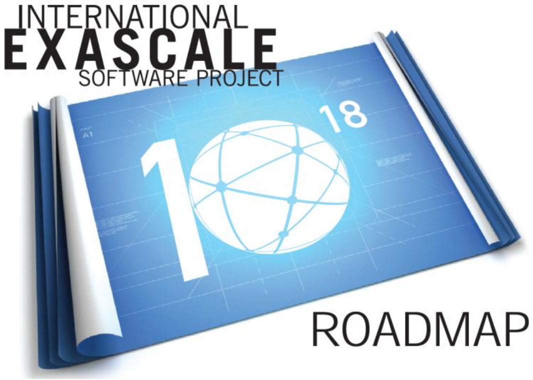
Application Efficiency: 🟨 System Libraries; 🟥 Computational Libraries; 🟨 Application Domain Libraries

Self-Aware – Security, Dependability, Efficiency

Some Predictions for DoE ExaScale Machines

Processors

- 10 billion-way concurrency
- 100's of cores per die
- 10 to 100-way per-core concurrency
- 100 million to 1 billion cores at 1 to 2 GHz
- Multi-threaded fine grain concurrency
- 10,000s of cycles system-wide latency

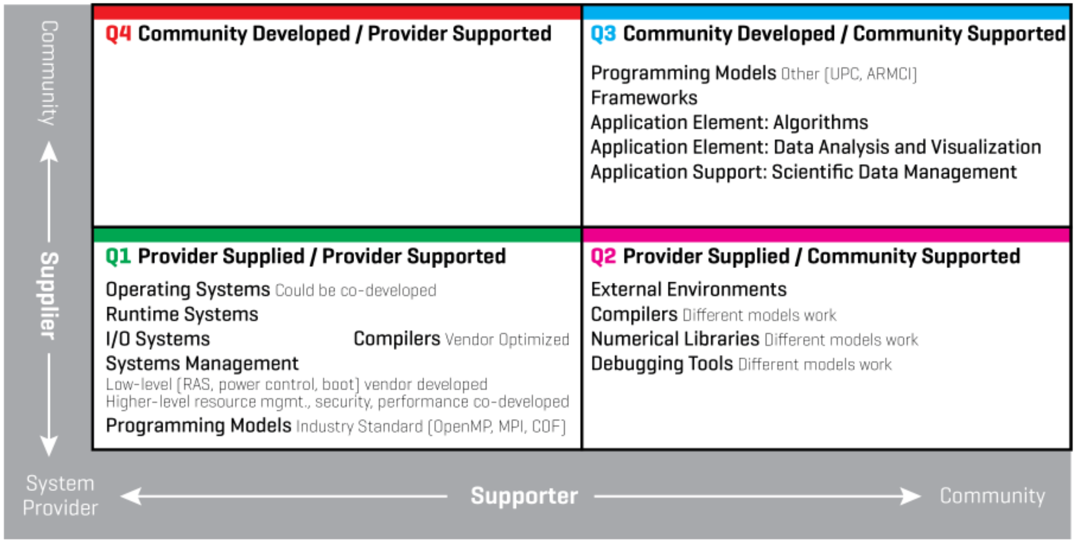


Memory

- Global address space without cache coherence
- Explicitly managed high speed buffer caches
- 128 PB capacity
- Deep memory hierarchies

Technology

- 22 to 11 nanometers CMOS
- 3-D packaging of dies
- Optical communications at 1TB/s
- Fault tolerance
- Active power management



Q4 Community Developed / Provider Supported

Q3 Community Developed / Community Supported

Programming Models Other (UPC, ARMCI)
Frameworks
Application Element: Algorithms
Application Element: Data Analysis and Visualization
Application Support: Scientific Data Management

Q1 Provider Supplied / Provider Supported

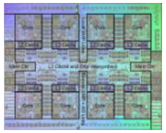
Operating Systems Could be co-developed
Runtime Systems
I/O Systems
Compilers Vendor Optimized
Systems Management
Low-level (RAS, power control, boot) vendor developed
Higher-level resource mgmt., security, performance co-developed
Programming Models Industry Standard (OpenMP, MPI, COF)

Q2 Provider Supplied / Community Supported

External Environments
Compilers Different models work
Numerical Libraries Different models work
Debugging Tools Different models work

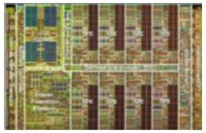
How Will They Look? Expect Evolution

- Multicore CPUs

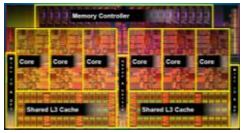


IBM POWER7
8 cores, 4-way SMT

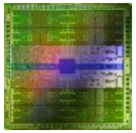
Intel SCC
48 cores



IBM Cell BE
8+1 cores



Intel Core i7
8 cores, 2-way SMT



Nvidia Fermi
448 cores, SMT

- Accelerators and distributed memory

RoadRunner
6,480 CPUs + 12,960 Cells
3240 TriBlades

HPC cabinet
CPU blades + GPU blades
Custom interconnect

Rack-mount server components
2 quad-core CPUs + 4 GPUs
200 Gflop/s + 4 Tflop/s

- Memory constrained

Dell PowerEdge R910
2 x 8-core CPUs
256 GB, 145 Gflop/s
1.7 byte/flop

BlueGene/L
65,536 dual-core CPUs
16 TB RAM, 360 Tflop/s
0.045 byte/flop

Nvidia Tesla M2050 (Fermi)
1 GPU, 448 cores
6 GB, 515 Gflop/s
0.011 byte/flop

HPC Software Development

Popular HPC programming languages

- 1953: Fortran
- 1973: C
- 1985: C++
- 1997: OpenMP
- 2007: CUDA

Popular HPC libraries

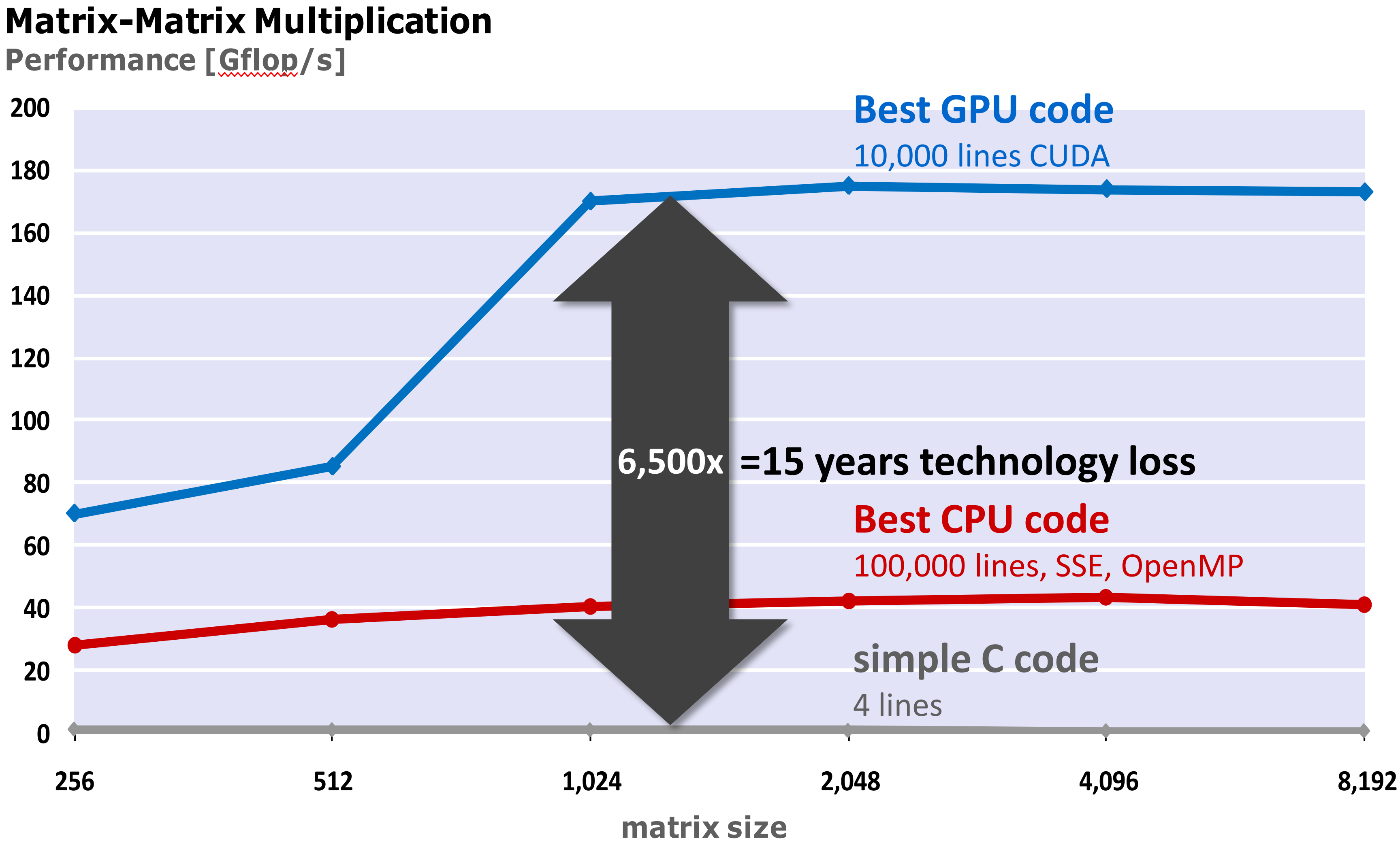
- 1979: BLAS
- 1992: LAPACK
- 1994: MPI
- 1995: ScaLAPACK
- 1995: PETSc
- 1997: FFTW

Proposed and maturing (?)

- Chapel, X10, Fortress, UPC, GA, HTA, OpenCL, Brook, Sequoia, Charm++, CnC, STAPL, TBB, Cilk,…

Slow change in direction

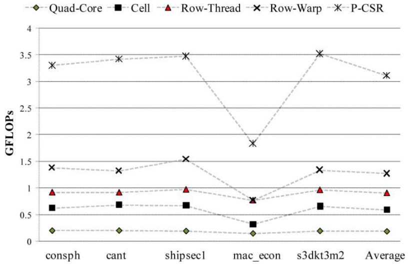
The Cost Of Portability and Maintainability

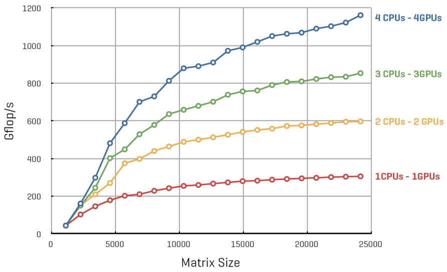


Summary

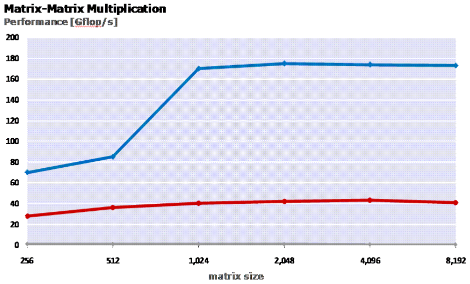
- **Hardware vendors will somehow keep Moore’s law on track**

- **Software optimization requires tuning experts with domain knowledge**





- **Portable and maintainable code costs (a lot of) performance**



Unoptimized program
= 15 years technology loss