2007 SPEC Benchmark Workshop
January 21, 2007
Radisson Hotel Austin North

The HPC Challenge Benchmark: A Candidate for Replacing LINPACK in the TOP500?

Jack Dongarra
University of Tennessee
and
Oak Ridge National Laboratory

Outline - The HPC Challenge Benchmark: A Candidate for Replacing Linpack in the TOP500?

- Look at LINPACK
- Brief discussion of DARPA HPCS Program
- HPC Challenge Benchmark
- Answer the Question

What Is LINPACK?

- Most people think LINPACK is a benchmark.
- LINPACK is a package of mathematical software for solving problems in linear algebra, mainly dense linear systems of linear equations.
- The project had its origins in 1974
- LINPACK: "LINear algebra PACKage"
  - Written in Fortran 66

Computing in 1974

- **High Performance Computers:**
  - IBM 370/195, CDC 7600, Univac 1110, DEC PDP-10, Honeywell 6030

- **Fortran 66**

- **Run efficiently**

• Inner product of 2 vectors	DOT	$w := \sum_{i=1}^{N} x_i y_i$
• Elementary vector operation	AXPY	$\mathbf{y} := a\mathbf{x} + \mathbf{y}$
• Givens plane rotation	ROTG/ROT	
• Modified Givens rotation	ROTMG/ROTM	
• Copy a vector **x** in **y**	COPY	$\mathbf{y} := \mathbf{x}$
• Interchange 2 vectors **x** and **y**	SWAP	$\mathbf{y} := \mathbf{x}$ and $\mathbf{x} := \mathbf{y}$
• Euclidian length (l_2-norm) of a vector	NRM 2	$w := \sqrt{\sum_{i=1}^{N} \mid x_i \mid^2}$
• Sum of absolute values of vector components	ASUM	$w := \sum_{i=1}^{N} \mid x_i \mid$
• Scaling of a vector	SCAL	$\mathbf{x} := a\mathbf{x}$
• Find largest component of a vector	AMAX	

- **BLAS (Level 1)**
  - Vector operations

- **Trying to achieve software portability**

- **LINPACK package was released in 1979**
  - About the time of the Cray 1

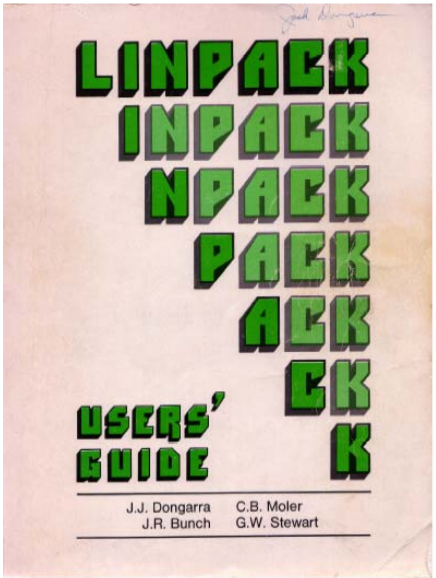
The Accidental Benchmarker

- **Appendix B of the Linpack Users' Guide**
  - Designed to help users extrapolate execution time for Linpack software package
- **First benchmark report from 1977;**
  - Cray 1 to DEC PDP-10

UNIT = 10**6 TIME/(1/3 100**3 + 100**2)

Facility	TIME N=100 secs.	UNIT micro-secs.	Computer	Type	Compiler
NCAR	.049	0.14	CRAY-1	S	CFT, Assembly BLAS
LASL	.148	0.43	CDC 7600	S	FTN, Assembly BLAS
NCAR	.192	0.56	CRAY-1	S	CFT
LASL	.210	0.61	CDC 7600	S	FTN
Argonne	.297	0.86	IBM 370/195	D	H
NCAR	.359	1.05	CDC 7600	S	Local
Argonne	.388	1.33	IBM 3033	D	H
NASA Langley	.489	1.42	CDC Cyber 175	S	FTN
U. Ill. Urbana	.506	1.47	CDC Cyber 175	S	Ext. 4.6
LLL	.554	1.61	CDC 7600	S	CHAT, No optimize
SLAC	.579	1.69	IBM 370/168	D	H Ext., Fast mult.
Michigan	.631	1.84	Amdahl 470/V6	D	H
Toronto	.890	2.59	IBM 370/165	D	H Ext., Fast mult.
Northwestern	1.44	4.20	CDC 6600	S	FTN
Texas	1.93*	5.63	CDC 6600	S	RUN
China Lake	1.95*	5.69	Univac 1110	S	V
Yale	2.59	7.53	DEC KL-20	S	F20
Bell Labs	3.46	10.1	Honeywell 6080	S	Y
Wisconsin	3.49	10.1	Univac 1110	S	V
Iowa State	3.54	10.2	Itel AS/5 mod3	D	H
U. Ill. Chicago	4.10	11.9	IBM 370/158	D	G1
Purdue	5.69	16.6	CDC 6500	S	FUN
U. C. San Diego	13.1	38.2	Burroughs 6700	S	H
Yale	17.1*	49.9	DEC KA-10	S	F40

* TIME(100) = (100/75)**3 SGEFA(75) + (100/75)**2 SGESL(75)



Dense matrices
Linear systems
Least squares problems
Singular values

LINPACK Benchmark?

- The LINPACK Benchmark is a measure of a computer's floating-point rate of execution for solving *Ax=b*.
  - It is determined by running a computer program that solves a dense system of linear equations.
- Information is collected and available in the LINPACK Benchmark Report.
- Over the years the characteristics of the benchmark has changed a bit.
  - In fact, there are three benchmarks included in the Linpack Benchmark report.
- LINPACK Benchmark since 1977
  - Dense linear system solve with LU factorization using partial pivoting
  - Operation count is: $2/3\ n^3 + O(n^2)$
  - Benchmark Measure: MFlop/s
  - Original benchmark measures the execution rate for a Fortran program on a matrix of size 100x100.

For Linpack with n = 100

- Not allowed to touch the code.
- Only set the optimization in the compiler and run.
- Provide historical look at computing
- Table 1 of the report (52 pages of 95 page report)
  - http://www.netlib.org/benchmark/performance.pdf

Computer	"LINPACK Benchmark" OS/Compiler	n=100 Mflop/s	"TPP" Best Effort n=1000 Mflop/s	"Theoritical Peak" Mflop/s
Intel Pentium Woodcrest (1 core, 3 GHz)	ifort -parallel -xT -O3 -ipo -mP2OPT_hlo_loop_unroll_factor=2	3018	6542	12000
Intel Pentium Woodcrest (1 core, 2.67 GHz)	ifort -O3 -ipo -xT -r8 -i8	2636		10680
NEC SX-8/8 (8proc. 2 GHz)			75140	128000
NEC SX-8/4 (4proc. 2 GHz)			43690	64000
NEC SX-8/2 (2proc. 2 GHz)			25060	32000
NEC SX-8/1 (1proc. 2 GHz)	-pi -Wf"-prob_use"	2177	14960	16000
HCL Infiniti Global Line 4700 HW (4 proc Intel Xeon 3.16 GHz)	ifort -fast -r8 -align	1892	9917	25280
HP ProLiant BL20p G3 (2 proc (1 cpu core per single chip), 3.8GHz Intel Xeon)			8185	14800

Linpack Benchmark Over Time

- In the beginning there was only the Linpack 100 Benchmark (1977)
  - n=100 (80KB); size that would fit in all the machines
  - Fortran; 64 bit floating point arithmetic
  - No hand optimization (only compiler options); source code available
- Linpack 1000 (1986)
  - n=1000 (8MB); wanted to see higher performance levels
  - Any language; 64 bit floating point arithmetic
  - Hand optimization OK
- Linpack Table 3 (Highly Parallel Computing - 1991) (Top500; 1993)
  - Any size (n as large as you can; n=10^6; 8TB; ~6 hours);
  - Any language; 64 bit floating point arithmetic
  - Hand optimization OK
    - Strassen's method not allowed (confuses the operation count and rate)
  - Reference implementation available
- In all cases results are verified by looking at: $\frac{\| Ax - b \|}{\| A \| \|x\| \, n \, \varepsilon} = O(1)$
- Operations count for factorization $\frac{2}{3}n^3 - \frac{1}{2}n^2$; solve $2n^2$

Motivation for Additional Benchmarks

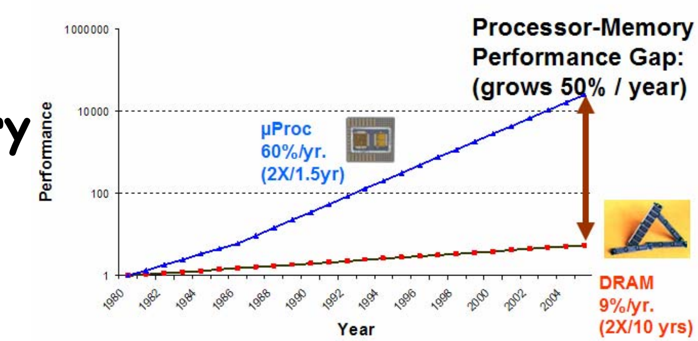
Linpack Benchmark

- Good
  - One number
  - Simple to define & easy to rank
  - Allows problem size to change with machine and over time
  - Stresses the system with a run of a few hours
- Bad
  - Emphasizes only "peak" CPU speed and number of CPUs
  - Does not stress local bandwidth
  - Does not stress the network
  - Does not test gather/scatter
  - Ignores Amdahl's Law (Only does weak scaling)
- Ugly
  - MachoFlops
  - Benchmarketeering hype

- From Linpack Benchmark and Top500: "no single number can reflect overall performance"
- Clearly need something more than Linpack
- HPC Challenge Benchmark
  - Test suite stresses not only the processors, but the memory system and the interconnect.
  - The real utility of the HPCC benchmarks are that architectures can be described with a wider range of metrics than just Flop/s from Linpack.

At The Time The Linpack Benchmark Was Created …

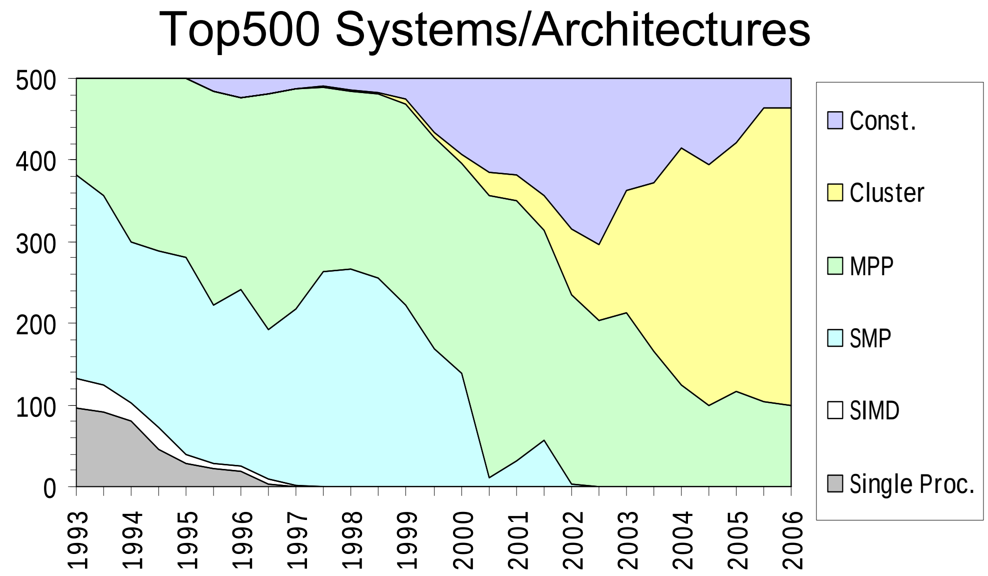
- If we think about computing in late 70's
- Perhaps the LINPACK benchmark was a reasonable thing to use.
- Memory wall, not so much a wall but a step.
- In the 70's, things were more in balance
  - The memory kept pace with the CPU
    - n cycles to execute an instruction, n cycles to bring in a word from memory



- Showed compiler optimization
- Today provides a historical base of data

Many Changes

- Many changes in our hardware over the past 30 years
  - Superscalar, Vector, Distributed Memory, Shared Memory, Multicore, …



- While there has been some changes to the Linpack Benchmark not all of them reflect the advances made in the hardware.

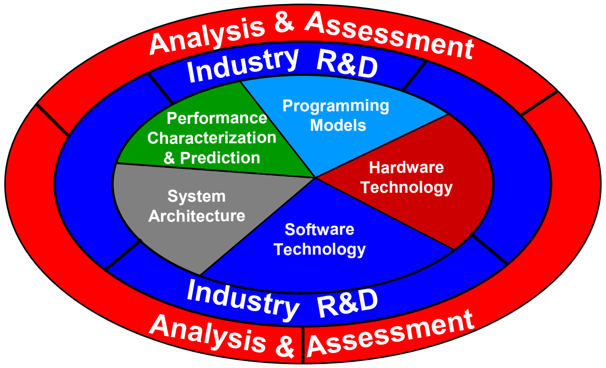
- Today's memory hierarchy is much more complicated.

High Productivity Computing Systems

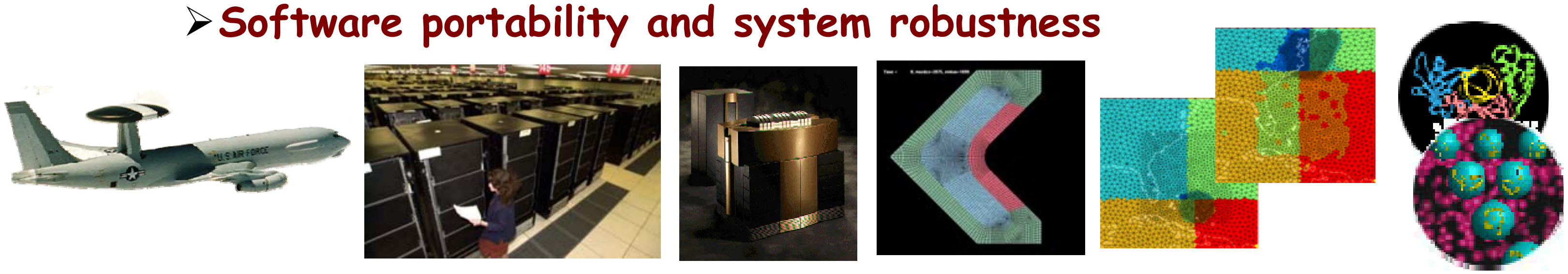
Goal:
Provide a generation of economically viable high productivity computing systems for the national security and industrial user community (2010; started in 2002)

Focus on:

- **Real (not peak) performance of critical national security applications**
  - Intelligence/surveillance
  - Reconnaissance
  - Cryptanalysis
  - Weapons analysis
  - Airborne contaminant modeling
  - Biotechnology
- **Programmability: reduce cost and time of developing applications**
- **Software portability and system robustness**



HPCS Program Focus Areas

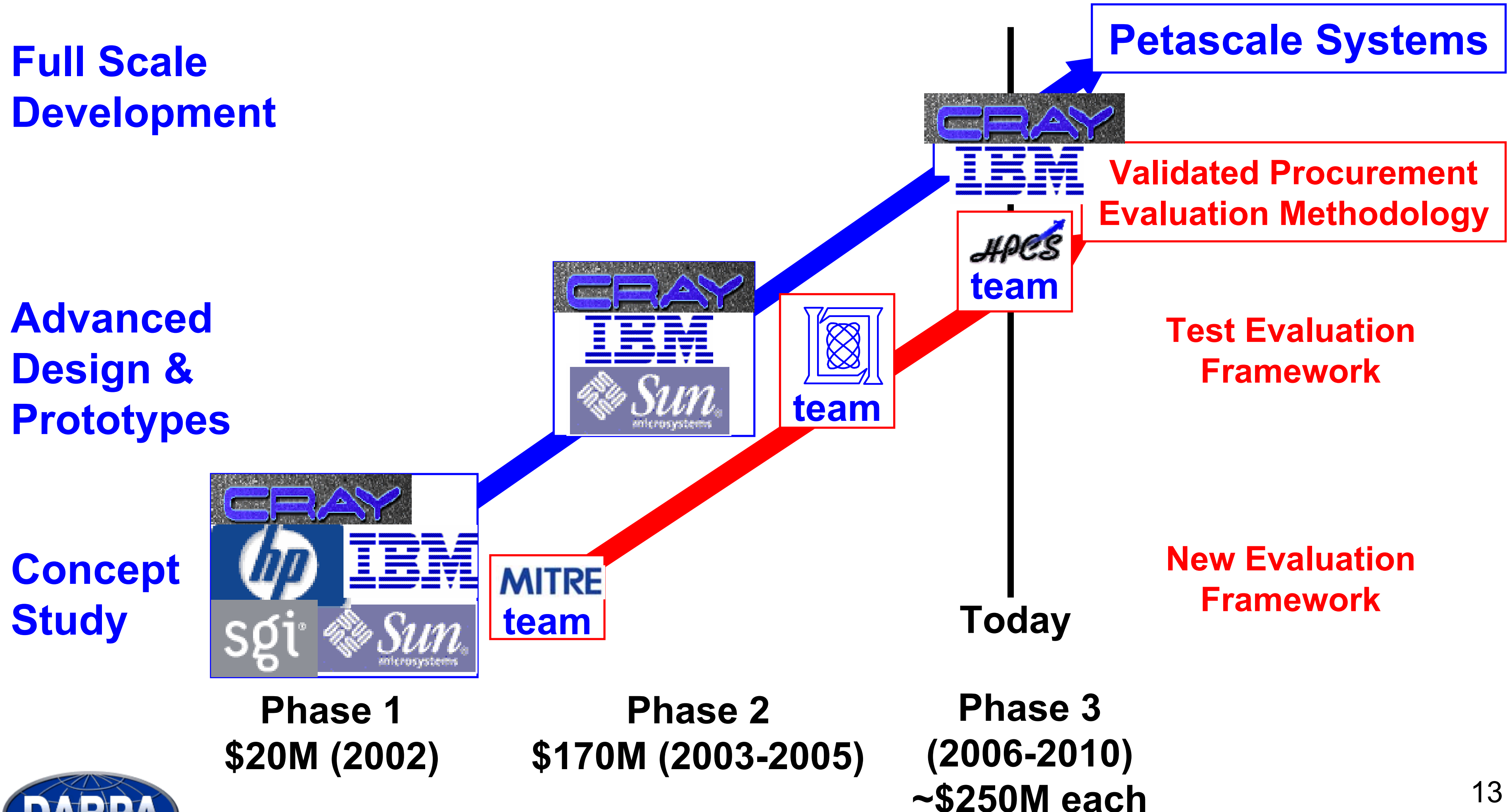


Applications:
Intelligence/surveillance, reconnaissance, cryptanalysis, weapons analysis, airborne contaminant modeling and biotechnology

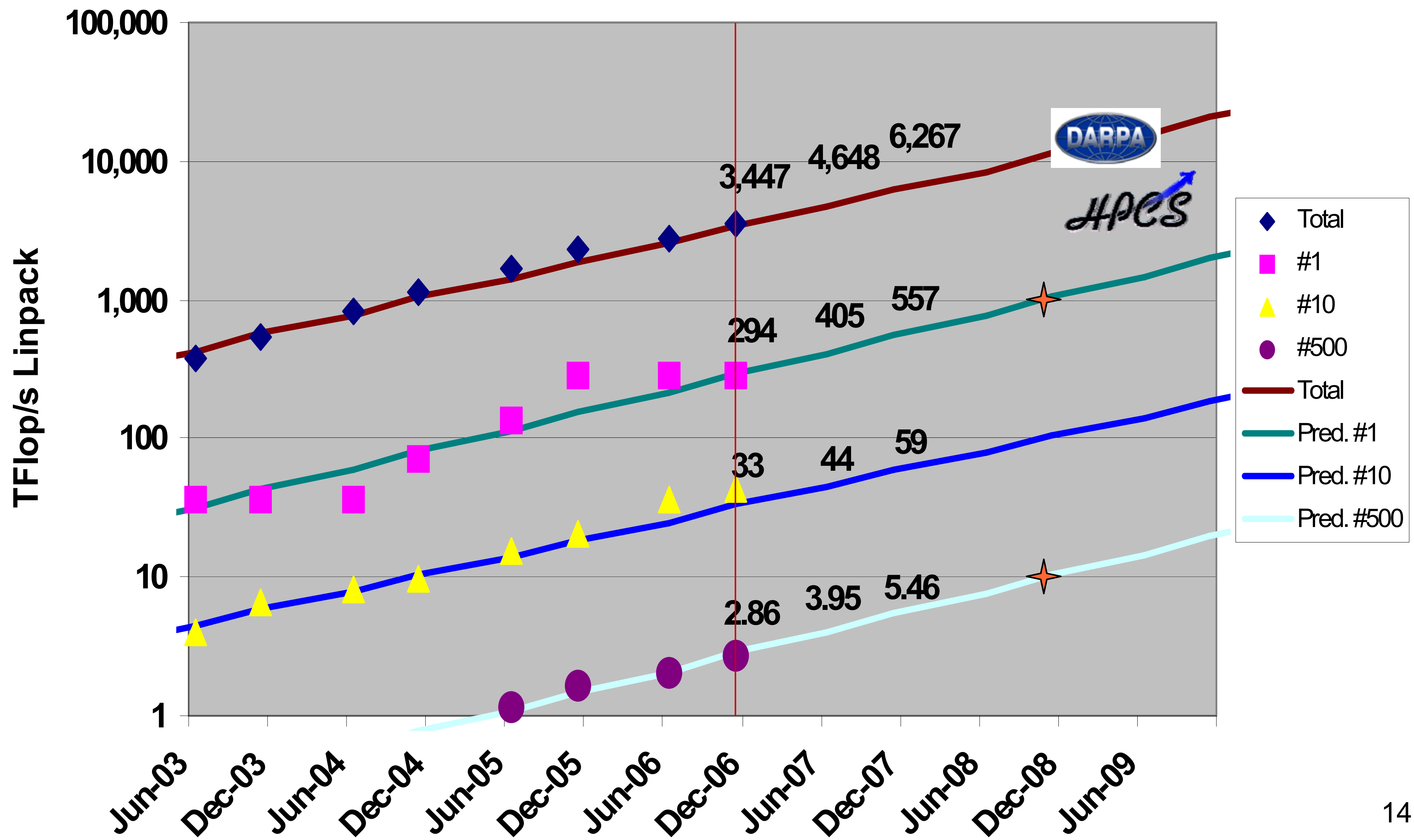
Fill the Critical Technology and Capability Gap
Today (late 80's HPC Technology) ... to ... Future (Quantum/Bio Computing)

HPCS Roadmap

- 5 vendors in phase 1; 3 vendors in phase 2; 1+ vendors in phase 3
- MIT Lincoln Laboratory leading measurement and evaluation team



Predicted Performance Levels for Top500



A PetaFlop Computer by the End of the Decade

- At least 10 Companies developing a Petaflop system in the next decade.

- Cray
- IBM
- Sun

} HPCS (DARPA): 2+ Pflop/s Linpack; 6.5 PB/s data streaming BW; 3.2 PB/s Bisection BW; 64,000 GUPS

- Dawning
- Galactic
- Lenovo

} Chinese Companies

- Hitachi
- NEC
- Fujitsu

} Japanese; "Life Simulator" (10 Pflop/s); Keisoku project $1B 7 years

- Bull

PetaFlop Computers in 2 Years!

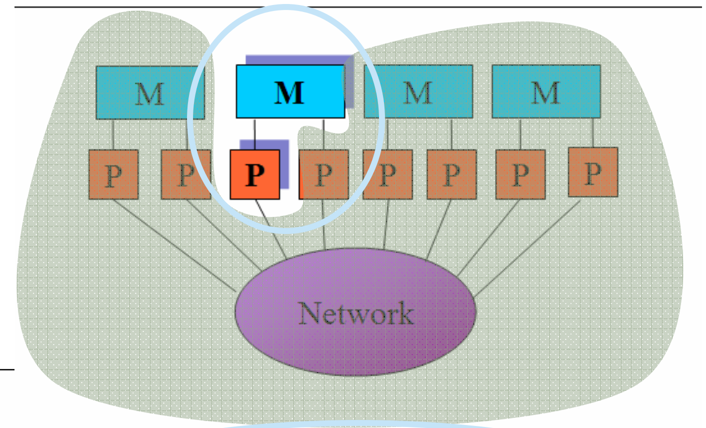
- Oak Ridge National Lab
  - Planned for 4th Quarter 2008 (1 Pflop/s peak)
  - From Cray's XT family
  - Use quad core from AMD
    - 23,936 Chips
    - Each chip is a quad core-processor (95,744 processors)
    - Each processor does 4 flops/cycle
    - Cycle time of 2.8 GHz
  - Hypercube connectivity
  - Interconnect based on Cray XT technology
  - 6MW, 136 cabinets
- Los Alamos National Lab
  - Roadrunner (2.4 Pflop/s peak)
  - Use IBM Cell and AMD processors
  - 75,000 cores

HPC Challenge Goals

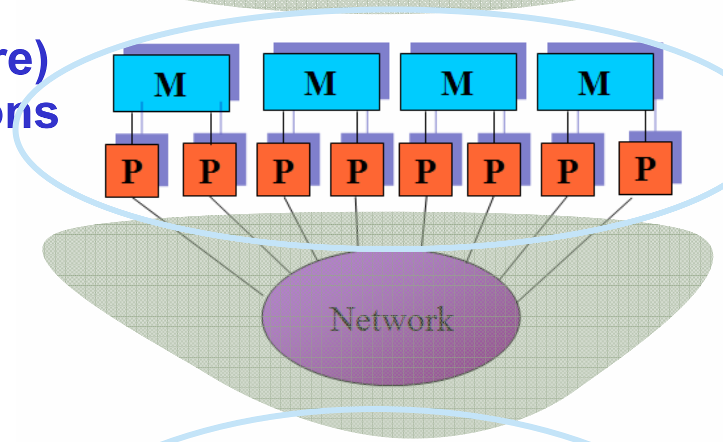
- To examine the performance of HPC architectures using kernels with more *challenging* memory access patterns than the Linpack Benchmark
  - The Linpack benchmark works well on all architectures — even cache-based, distributed memory multiprocessors due to
    1. Extensive memory reuse
    2. Scalable with respect to the amount of computation
    3. Scalable with respect to the communication volume
    4. Extensive optimization of the software
- To *complement* the Top500 list
- Stress CPU, memory system, interconnect
- Allow for optimizations
  - Record effort needed for tuning
  - Base run requires MPI and BLAS
- Provide verification & archiving of results

Tests on Single Processor and System

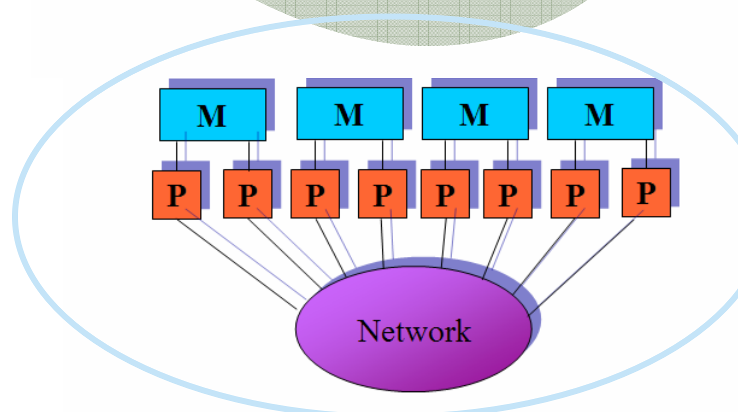
- **Local - only a single processor (core) is performing computations.**



- **Embarrassingly Parallel - each processor (core) in the entire system is performing computations but they do no communicate with each other explicitly.**



- **Global - all processors in the system are performing computations and they explicitly communicate with each other.**



HPC Challenge Benchmark

Consists of basically 7 benchmarks;

- Think of it as a framework or harness for adding benchmarks of interest.

1. LINPACK (HPL) — MPI Global ($Ax = b$)

2. STREAM — Local; single CPU
   *STREAM — Embarrassingly parallel

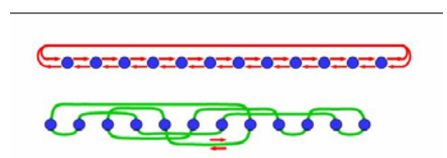
name	kernel	bytes/iter	FLOPS/iter
COPY:	a(i) = b(i)	16	0
SCALE:	a(i) = q*b(i)	16	1
SUM:	a(i) = b(i) + c(i)	24	1
TRIAD:	a(i) = b(i) + q*c(i)	24	2

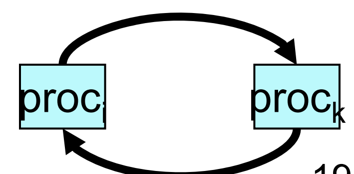
3. PTRANS ($A \leftarrow A + B^T$) — MPI Global

4. RandomAccess — Local; single CPU
   *RandomAccess — Embarrassingly parallel
   RandomAccess — MPI Global

Random integer read; update; & write

5. BW and Latency – MPI

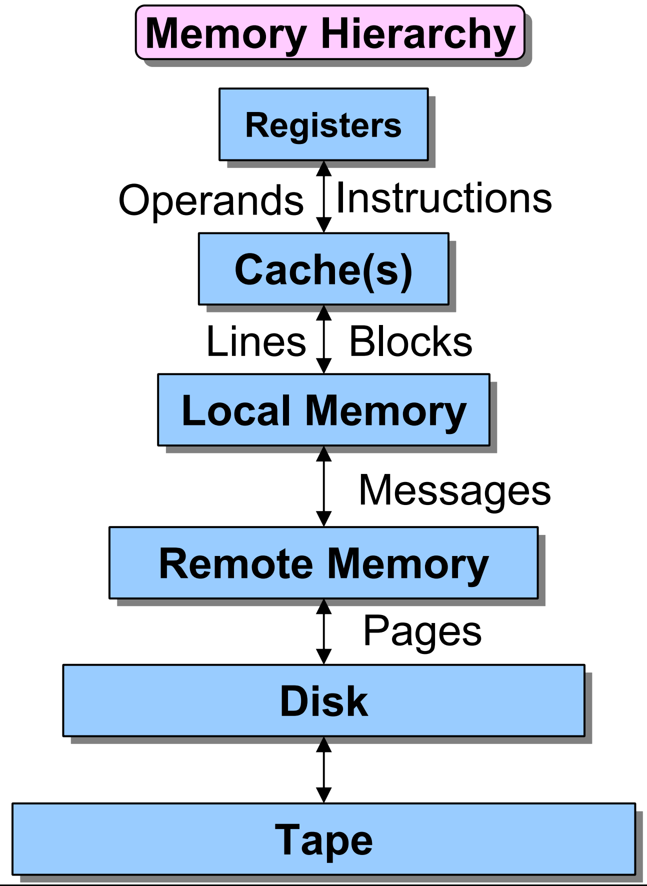




6. FFT - Global, single CPU, and EP

7. Matrix Multiply – single CPU and EP

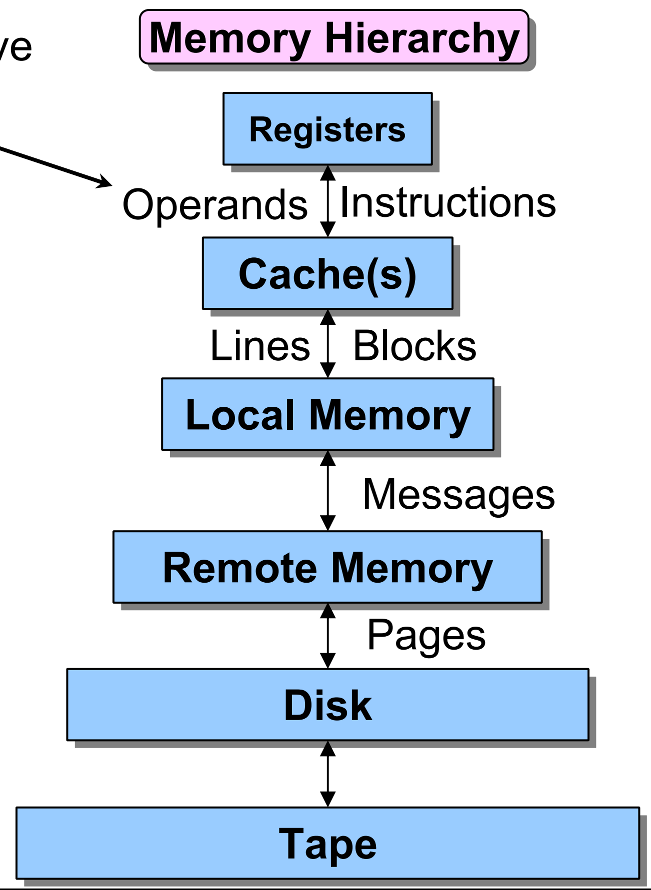
HPCS Performance Targets



- HPCC was developed by HPCS to assist in testing new HEC systems
- Each benchmark focuses on a different part of the memory hierarchy
- HPCS performance targets attempt to
  - Flatten the memory hierarchy
  - Improve real application performance
  - Make programming easier

HPCS Performance Targets

- LINPACK: linear system solve
  $Ax = b$

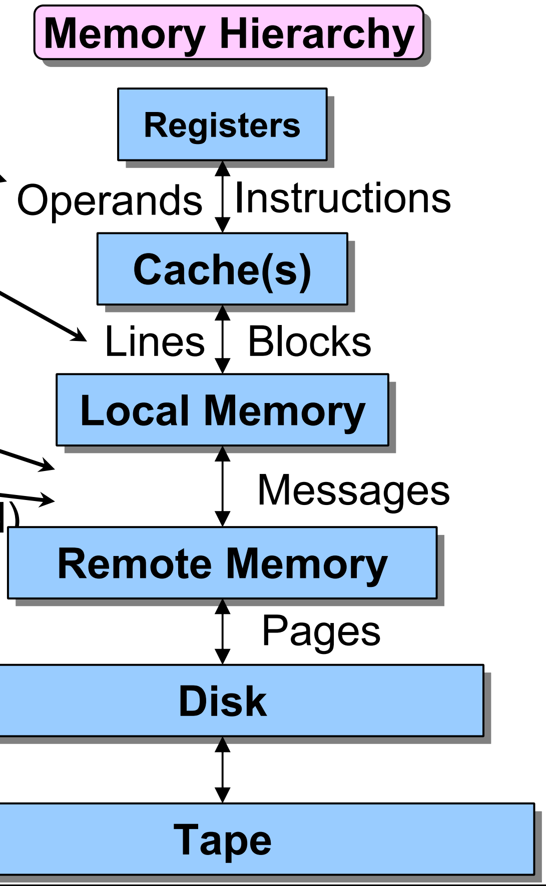


- **HPCC was developed by HPCS to assist in testing new HEC systems**
- **Each benchmark focuses on a different part of the memory hierarchy**
- **HPCS performance targets attempt to**
  - **Flatten the memory hierarchy**
  - **Improve real application performance**
  - **Make programming easier**

HPCS Performance Targets

- LINPACK: linear system solve
  $Ax = b$
- STREAM: vector operations
  $A = B + s * C$
- FFT: 1D Fast Fourier Transform
  $Z = fft(X)$
- RandomAccess: integer update
  $T[i] = XOR(T[i], rand)$

HPC Challenge



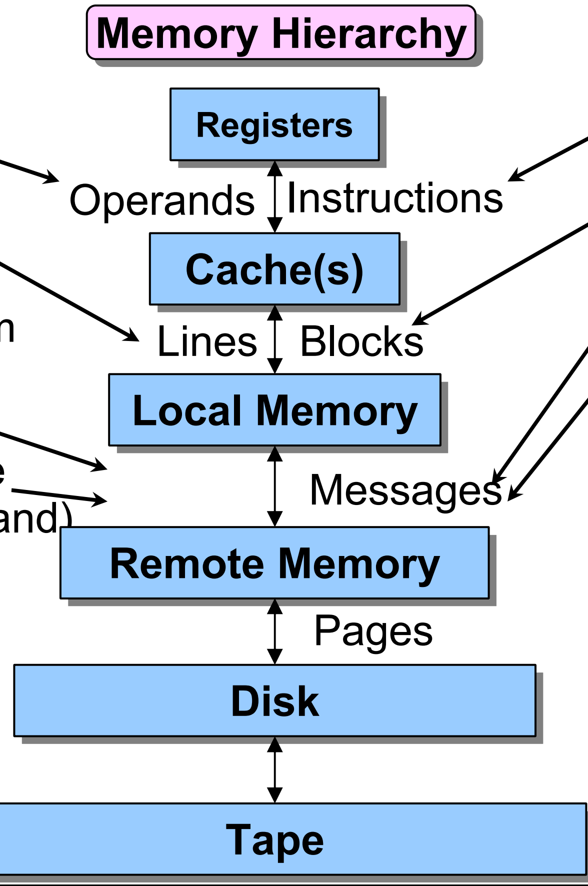
- **HPCC was developed by HPCS to assist in testing new HEC systems**
- **Each benchmark focuses on a different part of the memory hierarchy**
- **HPCS performance targets attempt to**
  - **Flatten the memory hierarchy**
  - **Improve real application performance**
  - **Make programming easier**

HPCS Performance Targets

- LINPACK: linear system solve
  Ax = b
- STREAM: vector operations
  A = B + s * C
- FFT: 1D Fast Fourier Transform
  Z = fft(X)
- RandomAccess: integer update
  T[i] = XOR(T[i], rand)

HPC Challenge

HPCC

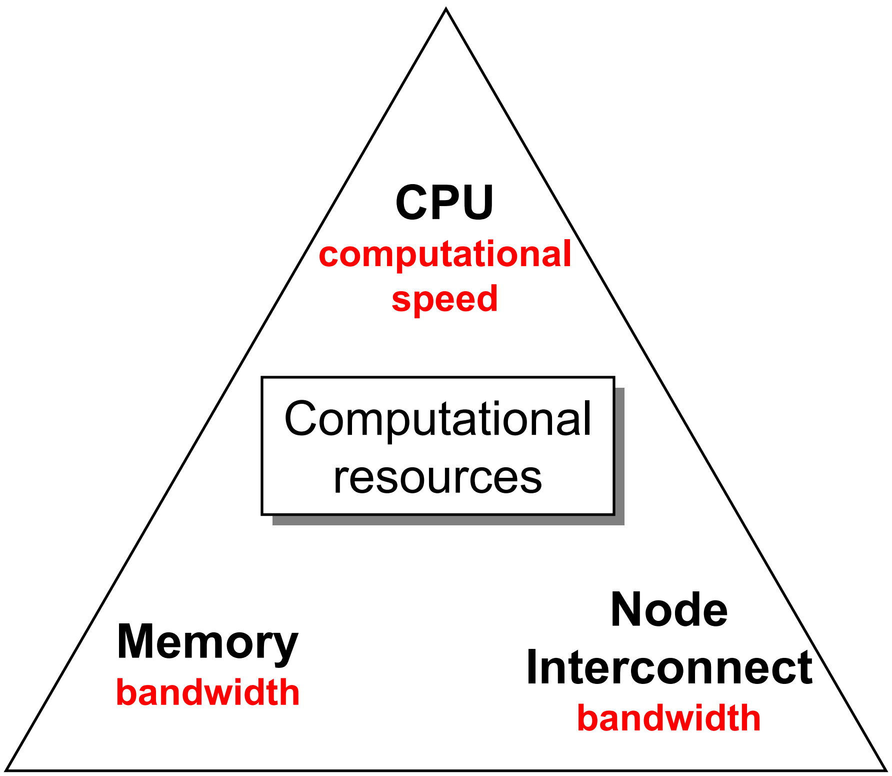


Max	Relative
2 Pflop/s	8x
6.5 Pbyte/s	40x
0.5 Pflop/s	200x
64000 GUPS	2000x

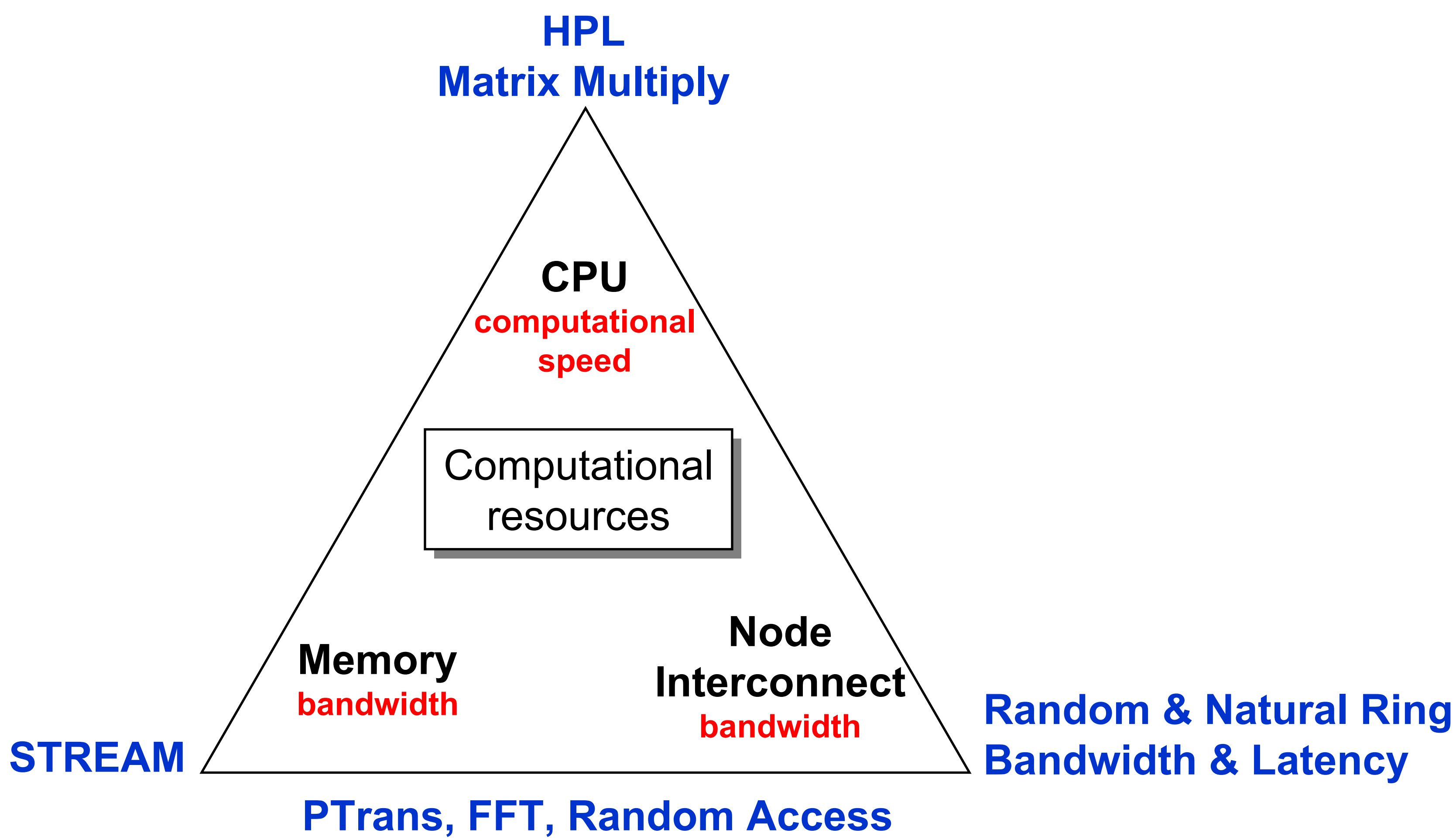
Performance Targets

- **HPCC was developed by HPCS to assist in testing new HEC systems**
- **Each benchmark focuses on a different part of the memory hierarchy**
- **HPCS performance targets attempt to**
  - **Flatten the memory hierarchy**
  - **Improve real application performance**
  - **Make programming easier**

Computational Resources and HPC Challenge Benchmarks



Computational Resources and HPC Challenge Benchmarks



How Does The Benchmarking Work?

- Single program to download and run
  - Simple input file similar to HPL input
- Base Run and Optimization Run
  - Base run must be made
    - User supplies MPI and the BLAS
  - Optimized run allowed to replace certain routines
    - User specifies what was done
- Results upload via website (monitored)
- html table and Excel spreadsheet generated with performance results
  - Intentionally we are not providing a single figure of merit (no over all ranking)
- Each run generates a record which contains 188 pieces of information from the benchmark run.
- Goal: no more than 2 X the time to execute HPL.

http://icl.cs.utk.edu/hpcc/ web

HPC CHALLENGE

HPCS

- Home
- Rules
- News
- Download
- FAQ
- Links
- Collaborators
- Sponsors
- Upload
- Kiviat Diagram
- Results

HPC Challenge Benchmark

The HPC Challenge benchmark consists of basically 7 benchmarks:

1. HPL - the Linpack TPP benchmark which measures the floating point rate of execution for solving a linear system of equations.
2. DGEMM - measures the floating point rate of execution of double precision real matrix-matrix multiplication.
3. STREAM - a simple synthetic benchmark program that measures sustainable memory bandwidth (in GB/s) and the corresponding computation rate for simple vector kernel.
4. PTRANS (parallel matrix transpose) - exercises the communications where pairs of processors communicate with each other simultaneously. It is a useful test of the total communications capacity of the network.
5. RandomAccess - measures the rate of integer random updates of memory (GUPS).
6. FFTE - measures the floating point rate of execution of double precision complex one-dimensional Discrete Fourier Transform (DFT).
7. Communication bandwidth and latency - a set of tests to measure latency and bandwidth of a number of simultaneous communication patterns; based on b_eff (effective bandwidth benchmark).

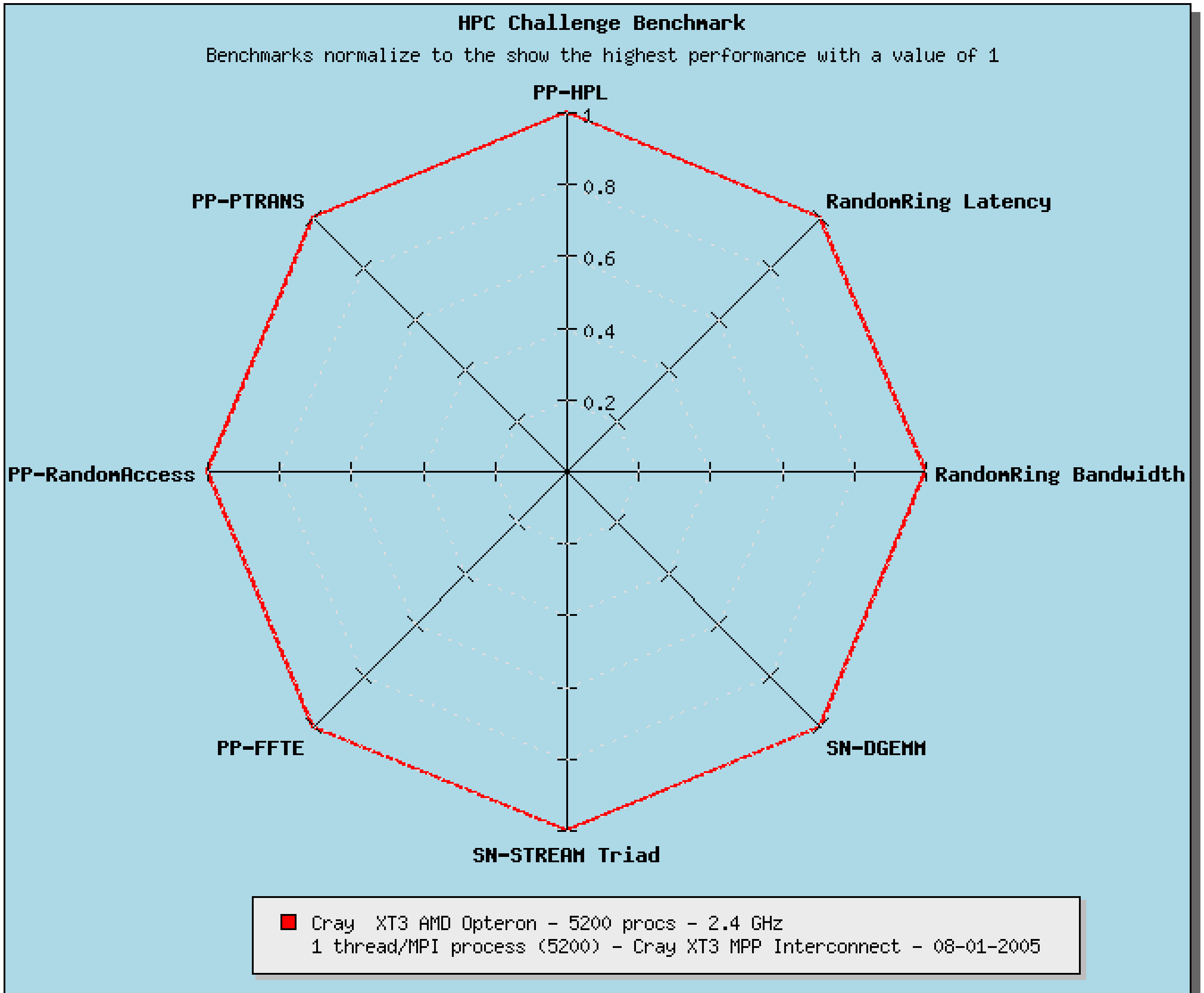
Condensed Results - Base Runs Only - 123 Systems - Generated on Sat Jan 13 14:12:19 2007

System Information System - Processor	Speed	Count	Threads	Processes	G-HPL	G-PTRANS	G-Random Access	G-FFTE	EP-STREAM Sys	EP-STREAM Triad	EP-DGEMM	RandomRing Bandwidth	RandomRing Latency
MA/PT/PS/PC/TH/PR/CM/CS/IC/IA/SD					TFlop/s	GB/s	Gup/s	GFlop/s	GB/s	GB/s	GFlop/s	GB/s	usec
Atipa Conquest cluster AMD Opteron	1.4GHz	128	1	128	0.2526110	3.2471			208.525	1.6291		0.03627	23.68
Clustervision BV Beastie AMD Opteron	2.4GHz	32	1	32	0.1037640	0.8159	0.0002350	2.1470	106.951	3.3422	4.19493	0.02648	53.23
Cray Inc. Red Storm/XT3 AMD Opteron	2.4GHz	12960	1	25920	91.0350000	2356.9700	1.7401500	1554.0700	54840.499	2.1158	4.39939	0.05911	16.29
Cray Inc. T3E Alpha 21164	0.6GHz	1024	1	1024	0.0481695	10.2765			529.242	0.5168		0.03174	12.09
Cray Inc. T3E Alpha 21164	0.675GHz	512	1	512	0.2231810	9.7741	0.0289464	15.4774	272.186	0.5316	0.66077	0.03571	8.14
Cray Inc. X1 Cray MSP	0.8GHz	64	1	64	0.5215600	3.2288			959.334	14.9896		0.94074	20.34
Cray Inc. X1 Cray MSP	0.8GHz	60	1	60	0.5777790	30.4313			898.446	14.9741		1.03291	20.83
Cray Inc. X1 Cray MSP	0.8GHz	120	1	120	1.0609700	2.4603			1019.519	8.4960		0.83014	20.12
Cray Inc. X1 Cray MSP	0.8GHz	252	1	252	2.3847300	97.4076			3758.404	14.9143		0.42899	22.27
Cray Inc. X1 Cray MSP	0.8GHz	124	1	124	1.2054200	39.5252			1856.664	14.9731		0.70857	20.15
Cray Inc. X1 Cray MSP	0.8GHz	60	1	60	0.5087430	1.6342	0.0030750	3.1444	894.114	14.9019	10.91520	1.16779	14.66
Cray Inc. X1 Cray MSP	0.8GHz	32	1	32	0.2767140	32.6606	0.0016620	2.9649	475.846	14.8702	8.25848	1.41269	14.94
Cray Inc. X1 Cray E	1.13GHz	1008	1	1008	12.0263000	108.0190	0.0861199	82.3884	15522.091	15.3989	14.50000	0.15667	16.30
Cray Inc. X1E Cray X1 MSP	1.13GHz	252	1	252	3.1940900	85.2040	0.0148684	15.5352	2439.985	9.6825	14.18470	0.36024	14.93
Cray Inc. X1E Cray	1.13GHz	32	4	32	0.3376360	18.9199	0.0089686	5.2027	307.565	9.6114	11.60560	1.40487	12.21
Cray Inc. XD1 AMD Opteron	2.2GHz	64	1	64	0.2238980	10.5924	0.0223966	16.3611	169.955	2.6555	4.03375	0.22697	1.63
System Information System - Processor	Speed	Count	Threads	Processes	G-HPL	G-PTRANS	G-Random Access	G-FFTE	EP-STREAM Sys	EP-STREAM Triad	EP-DGEMM	RandomRing Bandwidth	RandomRing Latency
MA/PT/PS/PC/TH/PR/CM/CS/IC/IA/SD					TFlop/s	GB/s	Gup/s	GFlop/s	GB/s	GB/s	GFlop/s	GB/s	usec
Cray Inc. XD1 AMD Opteron	2.4GHz	128	1	128	0.5020760	13.5155	0.0666722	35.5172	500.065	3.9068	4.33435	0.25919	2.06
Cray Inc. XT3 AMD Opteron	2.6GHz	1100	1	1100	4.7823400	217.9230	0.1370020	266.6600	5274.698	4.7952	4.81050	0.28638	25.94
Cray Inc. XT3 AMD Opteron	2.4GHz	3744	1	3744	14.7040000	608.5060	0.2202960	417.1720	18146.382	4.8468	4.41330	0.16164	25.32
Cray Inc. XT3 AMD Opteron	2.4GHz	5200	1	5200	20.5270000	874.8990	0.2685830	644.7300	26020.800	5.0040	4.39535	0.14682	25.80
Cray Inc. xt3 AMD Opteron	2.4GHz	32	1	32	0.1387810	7.3764	0.0606017	9.3683	156.424	4.8883	4.77641	0.57281	8.74
Cray Inc. XT3 AMD Opteron	2.6GHz	4096	1	4096	16.9752000	302.9790	0.5330720	905.5690	20656.456	5.0431	4.78166	0.16896	9.44

Condensed Results - Base and Optimized Runs - 144 Systems - Generated on Sat Jan 13 14:14:56 2007

System Information System - Processor - Speed - Count - Threads - Processes	Run	G-HPL	G-PTRANS	G-Random Access	G-FFTE	EP-STREAM Sys	EP-STREAM Triad	EP-DGEMM	RandomRing Bandwidth	RandomRing Latency
MA/PT/PS/PC/TH/PR/CM/CS/IC/IA/SD	Type	TFlop/s	GB/s	Gup/s	GFlop/s	GB/s	GB/s	GFlop/s	GB/s	usec
IBM Blue Gene/L PowerPC 440 0.7GHz 131072 165536	opt	**259.2130000**	374.4180	32.9834000	2228.39	159898.665	2.4399	2.31471	0.01110	7.78
IBM Blue Gene/L PowerPC 440 0.7GHz 131072 165536	opt	252.2970000	369.6300	**35.4706000**	**2311.09**	**160064.471**	2.4424	2.07220	0.01105	7.89
Cray Inc. Red Storm/XT3 AMD Opteron 2.4GHz 12960 125920	base	91.0350000	2356.9700	1.7401500	1554.07	54840.499	2.1158	4.39939	0.05911	16.29
Cray Inc. Red Storm/XT3 AMD Opteron 2.4GHz 12960 125920	opt	90.9902000	2351.5000	29.8180000	1529.14	53891.827	2.0792	4.39936	0.05910	15.76
IBM Blue Gene/L PowerPC 440 0.7GHz 65536 165536	base	80.6830000	339.2840	0.0657312	2178.11	53555.888	0.8172	1.85619	0.01084	8.84
IBM Blue Gene/L PowerPC 440 0.7GHz 32768 116384	opt	**67.1174000**	137.2380	**17.2911000**	988.18	39984.169	2.4404	2.31468	0.02186	5.88
IBM p5-575 Power5 1.9GHz 10240 110240	base	**57.8670000**	553.0090	0.1693440	842.50	**55184.179**	5.3891	7.08562	0.11015	118.59
IBM p5-575 Power5 1.9GHz 8192 18192	base	45.7019000	2626.1700	0.3239760	908.69	44455.936	5.4268	7.06423	0.08871	11.05
Cray Inc. XT3 Dual-Core AMD Opteron 2.6GHz 10404 110404	opt	43.5056000	2038.9200	**10.6711000**	**1122.70**	26539.772	2.5509	4.79356	0.08201	17.04
Cray Inc. XT3 Dual-Core AMD Opteron 2.6GHz 10404 110404	base	43.4033000	778.3850	0.8235630	1107.21	25774.557	2.4774	4.78995	0.06937	14.32
IBM Blue Gene/L PowerPC 440 0.7GHz 65536 165536	base	37.3540000	4665.9100	0.1648600	1762.82	62889.787	0.9596	2.47017	0.01039	8.62
IBM p5-575 Power5 1.9GHz 8192 28192	base	33.3175000	575.8230	0.2066390	966.67	43802.460	5.3470	6.08616	0.07698	51.99
Cray Inc. XT3 AMD Opteron 2GHz 10350 110350	base	32.9865000	1813.0600	1.0176500	**1118.29**	**43581.780**	4.2108	3.66719	0.16188	10.32
IBM Blue Gene/L PowerPC 440 0.7GHz 32768 132768	base	31.2581000	87.7818	0.2780090	1112.81	29913.678	0.9129	2.17447	0.01197	9.51
Cray Inc. XT3 AMD Opteron 2.4GHz 5200 15200	base	20.5270000	874.8990	0.2685830	644.73	26020.800	5.0040	4.39535	0.14682	25.80
Cray Inc. XT3 AMD Opteron 2.4GHz 5208 15208	opt	20.4163000	942.2520	0.6600460	779.43	29318.540	5.6295	4.41290	0.20474	9.33
System Information System - Processor - Speed - Count - Threads - Processes	Run	G-HPL	G-PTRANS	G-Random Access	G-FFTE	EP-STREAM Sys	EP-STREAM Triad	EP-DGEMM	RandomRing Bandwidth	RandomRing Latency
MA/PT/PS/PC/TH/PR/CM/CS/IC/IA/SD	Type	TFlop/s	GB/s	Gup/s	GFlop/s	GB/s	GB/s	GFlop/s	GB/s	usec
Cray Inc. XT3 AMD Opteron 2.4GHz 5208 15208	opt	20.4163000	942.2520	0.6600460	779.43	29318.540	5.6295	4.41290	0.20474	9.33
Cray Inc. XT3 AMD Opteron 2.4GHz 5208 15208	base	20.4086000	944.2270	0.6724120	761.73	24268.447	4.6598	4.41173	0.20636	9.20
Cray Inc. XT3 AMD Opteron 2.4GHz 5208 15208	opt	20.3371000	944.2090	0.6874420	855.24	29218.494	5.6103	4.41835	0.19878	9.18
Cray Inc. XT3 AMD Opteron 2.6GHz 4096 14096	base	16.9752000	302.9790	0.5330720	905.57	20656.456	5.0431	4.78166	0.16896	9.44
Cray Inc. XT3 AMD Opteron 2.6GHz 4128 14128	base	16.6421000	674.7860	0.6767580	821.68	19295.676	4.6743	4.75946	0.22245	8.23
Cray Inc. XT3 AMD Opteron 2.4GHz 3744 13744	base	14.7040000	608.5060	0.2202960	417.17	18146.382	4.8468	4.41330	0.16164	25.32
Cray Inc. X1 Cray E 1.13GHz 1008 11008	opt	12.2650000	144.9730	7.6881900	245.09	12687.293	12.5866	14.17580	0.15317	16.30

HPCC Kiviat Chart

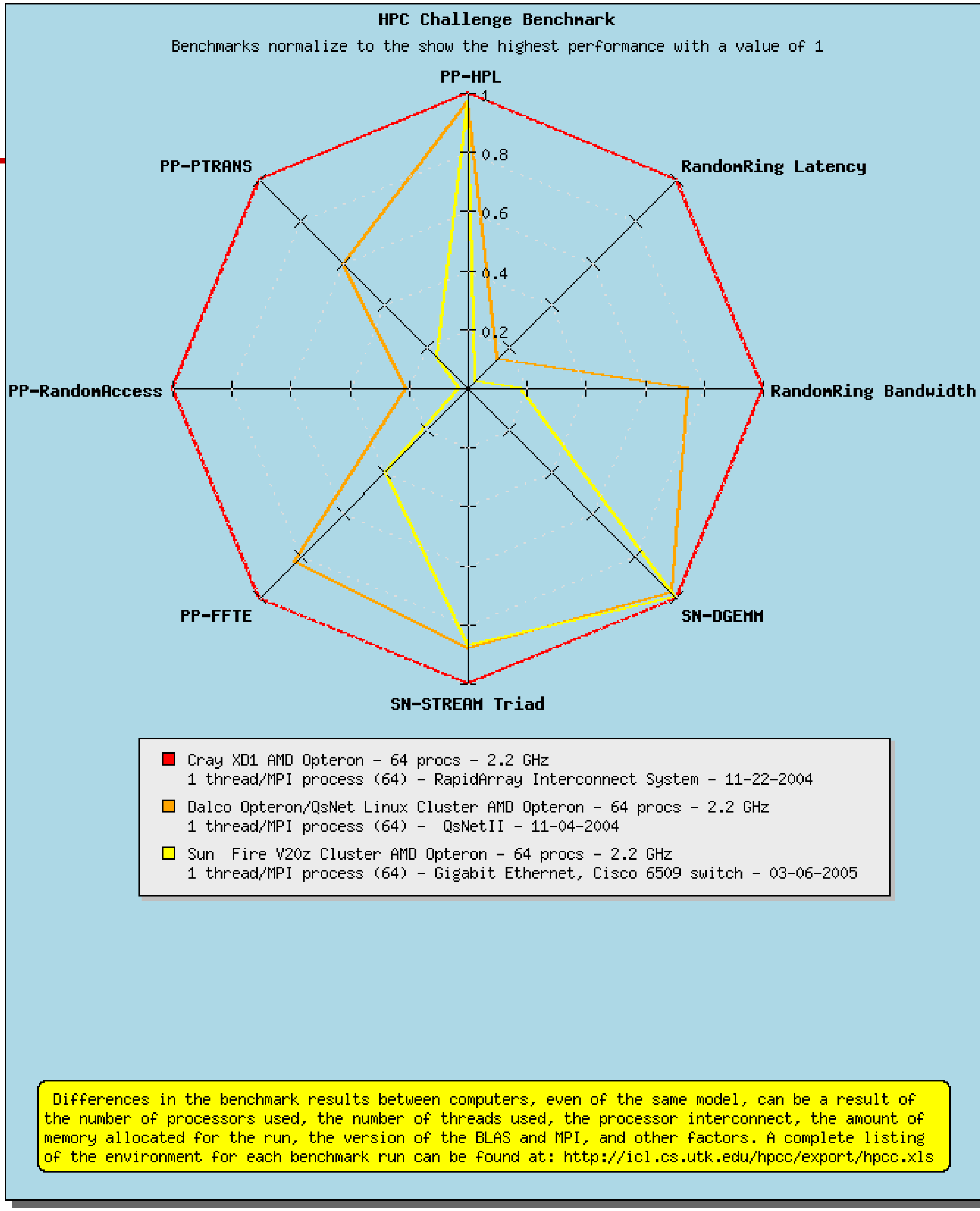


http://icl.cs.utk.edu/hpcc/

The values plotted for HPL, PTRANS, RandomAccess, and FFTE are per processor. The values plotted for SN-DGEMM and SN-STREAM are per thread. The value plotted for RandomRing Latency is normalized using it's reciprocal. Only those systems that have values for all the tests plotted are available for the diagram. Use the left-hand column to select up to 6 systems to plot in the Kiviat diagram.

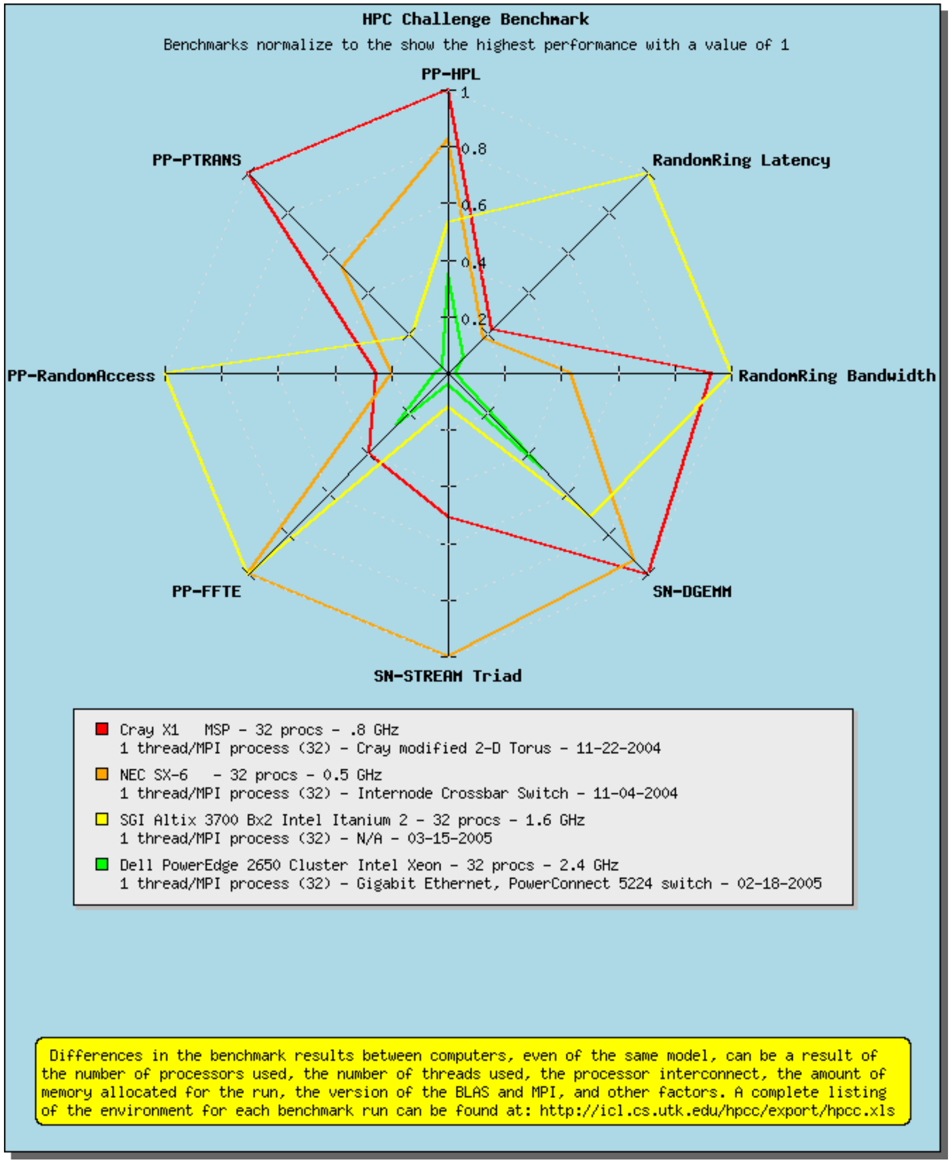
Systems for Kiviat Chart - Base Runs Only - 100 Systems - Generated on Sat Jan 13 14:16:51 2007

Plot	System Information System - Processor - Speed - Count - Threads - Processes	PP-HPL	PP-PTRANS	PP-Random Access	PT-SN-STREAM Triad	PP-FFTE	PT-SN-DGEMM	RandomRing Bandwidth	RandomRing Latency
	MA/PT/PS/PC/TH/PR/CM/CS/IC/IA/SD	TFlop/s	GB/s	Gup/s	GB/s	GFlop/s	GFlop/s	GB/s	usec
☐	Clustervision BV Beastie AMD Opteron 2.4GHz 32 1 32	0.00324262	0.025498	(0.00000734)	3.3391	(0.067094)	4.19992	0.02648	53.23
☐	Cray Inc. Red Storm/XT3 AMD Opteron 2.4GHz 12960 125920	0.00702431	0.181865	(0.00013427)	4.0983	(0.119913)	4.40779	0.05911	16.29
☐	Cray Inc. T3E Alpha 21164 0.675GHz 512 1 512	0.00043590	0.019090	(0.00005654)	0.5422	(0.030229)	0.68034	0.03571	8.14
☐	Cray Inc. X1 Cray MSP 0.8GHz 60 1 60	0.00847905	0.027237	(0.00005125)	16.2112	(0.052406)	10.90440	1.16779	14.66
☐	Cray Inc. X1 Cray MSP 0.8GHz 32 1 32	0.00864731	1.020644	(0.00005194)	16.2214	(0.092654)	8.45943	1.41269	14.94
☐	Cray Inc. X1 Cray E 1.13GHz 1008 11008	0.01193085	0.107162	(0.00008544)	32.7060	(0.081735)	15.25730	0.15667	16.30
☐	Cray Inc. X1E Cray X1 MSP 1.13GHz 252 1 252	0.01267496	0.338111	(0.00005900)	23.1291	(0.061648)	15.15610	0.36024	14.93
☐	Cray Inc. X1E Cray 1.13GHz 32 4 32	0.01055112	0.591247	(0.00028027)	5.7105	(0.162583)	3.62873	1.40487	12.21
☐	Cray Inc. XD1 AMD Opteron 2.2GHz 64 1 64	0.00349841	0.165506	(0.00034995)	2.7662	(0.255642)	3.98010	0.22697	1.63
☐	Cray Inc. XD1 AMD Opteron 2.4GHz 128 1 128	0.00392247	0.105590	(0.00052088)	4.3576	(0.277478)	4.33436	0.25919	2.06
☐	Cray Inc. XT3 AMD Opteron 2.6GHz 1100 11100	0.00434758	0.198112	(0.00012455)	4.9892	(0.242418)	4.81096	0.28638	25.94
☐	Cray Inc. XT3 AMD Opteron 2.4GHz 3744 13744	0.00392735	0.162528	(0.00005884)	4.6212	(0.111424)	4.41419	0.16164	25.32
☐	Cray Inc. XT3 AMD Opteron 2.4GHz 5200 15200	0.00394750	0.168250	(0.00005165)	4.7202	(0.123987)	4.39289	0.14682	25.80
☐	Cray Inc. xt3 AMD Opteron 2.4GHz 32 1 32	0.00433691	0.230513	(0.00189380)	4.8882	(0.292758)	4.77263	0.57281	8.74
☐	Cray Inc. XT3 AMD Opteron 2.6GHz 4096 14096	0.00414434	0.073969	(0.00013014)	5.0423	(0.221086)	4.77510	0.16896	9.44
☐	Cray Inc. XT3 AMD Opteron 2.4GHz 5208 15208	0.00391870	0.181303	(0.00012911)	4.6028	(0.146261)	4.41321	0.20636	9.20
Plot	System Information System - Processor - Speed - Count - Threads - Processes	PP-HPL	PP-PTRANS	PP-Random Access	PT-SN-STREAM Triad	PP-FFTE	PT-SN-DGEMM	RandomRing Bandwidth	RandomRing Latency
	MA/PT/PS/PC/TH/PR/CM/CS/IC/IA/SD	TFlop/s	GB/s	Gup/s	GB/s	GFlop/s	GFlop/s	GB/s	usec
☐	Cray Inc. XT3 AMD Opteron 2.6GHz 4128 14128	0.00403152	0.163466	(0.00016394)	4.6202	(0.199050)	4.75482	0.22245	8.23
☐	Cray Inc. XT3 AMD Opteron 2GHz 10350 110350	0.00318710	0.175175	(0.00009832)	4.3689	(0.108047)	3.67306	0.16188	10.32
☐	Cray Inc. XT3 AMD Opteron 2.6GHz 1100 11100	0.00429787	0.230315	(0.00027597)	4.8756	(0.298442)	4.77169	0.39964	7.29



HPC Challenge Benchmark
Benchmarks normalize to the show the highest performance with a value of 1
PP-HPL
RandomRing Latency
RandomRing Bandwidth
SN-DGEMM
SN-STREAM Triad
PP-FFTE
PP-RandomAccess
PP-PTRANS
1
0.8
0.6
0.4
0.2
Cray XD1 AMD Opteron - 64 procs - 2.2 GHz
1 thread/MPI process (64) - RapidArray Interconnect System - 11-22-2004
Dalco Opteron/QsNet Linux Cluster AMD Opteron - 64 procs - 2.2 GHz
1 thread/MPI process (64) - QsNetII - 11-04-2004
Sun Fire V20z Cluster AMD Opteron - 64 procs - 2.2 GHz
1 thread/MPI process (64) - Gigabit Ethernet, Cisco 6509 switch - 03-06-2005
Differences in the benchmark results between computers, even of the same model, can be a result of the number of processors used, the number of threads used, the processor interconnect, the amount of memory allocated for the run, the version of the BLAS and MPI, and other factors. A complete listing of the environment for each benchmark run can be found at: http://icl.cs.utk.edu/hpcc/export/hpcc.xls

Different Computers are Better at Different Things, No "Fastest" Computer for All Aps



HPCC Awards Info and Rules

Class 1 (Objective)

- **Performance**
  1. G-HPL $500
  2. G-RandomAccess $500
  3. EP-STREAM system $500
  4. G-FFT $500
- **Must be full submissions through the HPCC database**

Winners (in both classes) will be announced at SC07 HPCC BOF

Sponsored by:

Class 2 (Subjective)

- **Productivity (Elegant Implementation)**
  - Implement at least two tests from Class 1
  - $1500 (may be split)
  - Deadline:
    - October 15, 2007
  - Select 3 as finalists
- **This award is weighted**
  - 50% on performance and
  - 50% on code elegance, clarity, and size.
- **Submissions format flexible**

Class 2 Awards

- Subjective
- Productivity (Elegant Implementation)
  - Implement at least two tests from Class 1
  - $1500 (may be split)
  - Deadline:
    - October 15, 2007
  - Select 5 as finalists
- Most "elegant" implementation with special emphasis being placed on:
- Global HPL, Global RandomAccess, EP STREAM (Triad) per system and Global FFT.
- This award is weighted
  - 50% on performance and
  - 50% on code elegance, clarity, and size.

5 Finalists for Class 2 – November 2005

- Cleve Moler, Mathworks
  - Environment: Parallel Matlab Prototype
  - System: 4 Processor Opteron
- Calin Caseval, C. Bartin, G. Almasi, Y. Zheng, M. Farreras, P. Luk, and R. Mak, IBM
  - Environment: UPC
  - System: Blue Gene L
- Bradley Kuszmaul, MIT
  - Environment: Cilk
  - System: 4-processor 1.4Ghz AMD Opteron 840 with 16GiB of memory
- Nathan Wichman, Cray
  - Environment: UPC
  - System: Cray X1E (ORNL)
- Petr Konency, Simon Kahan, and John Feo, Cray
  - Environment: C + MTA pragmas
  - System: Cray MTA2

Winners!

2006 Competitors

- **Some Notable Class 1 Competitors**

Cray (DOD ERDC)
XT3 4096 CPUs
Sapphire

SGI (NASA)
Columbia
10,000 CPUs

NEC (HLRS)
SX-8 512 CPUs

Cray
(Sandia)
XT3 11,648 CPU
Red Storm

IBM (DOE LLNL)
BG/L 131,072 CPU
Purple 10,240 CPU

Cray (DOE ORNL)
X1 1008 CPUs
Jaguar XT3 5200 CPUs

DELL (MIT LL)
300 CPUs
LLGrid

- **Class 2: 6 Finalists**
  - Calin Cascaval (IBM) UPC on Blue Gene/L [Current Language]
  - Bradley Kuszmaul (MIT CSAIL) Cilk on SGI Altix [Current Language]
  - Cleve Moler (Mathworks) Parallel Matlab on a cluster [Current Language]
  - Brad Chamberlain (Cray) Chapel [Research Language]
  - Vivek Sarkar (IBM) X10 [Research Language]
  - Vadim Gurev (St. Petersburg, Russia) MCSharp [Student Submission]

The Following are the Winners of the 2006 HPC Challenge Class 1 Awards

G-HPL	Achieved	System	Affiliation	Submitter
1st place	259 Tflop/s	IBM BG/L	DOE/NNSA/LLNL	Tom Spelce
1st runner up	67 Tflop/s	IBM BG/L	IBM T.J. Watson	John Gunnels
2nd runner up	57 Tflop/s	IBM p5-575	LLNL	Charles Grassl

G-RandomAccess	Achieved	System	Affiliation	Submitter
1st place	35 GUPS	IBM BG/L	DOE/NNSA/LLNL	Tom Spelce
1st runner up	17 GUPS	IBM BG/L	IBM T.J. Watson	John Gunnels
2nd runner up	10 GUPS	Cray XT3 Dual	ORNL	Jeff Larkin

G-FFT	Achieved	System	Affiliation	Submitter
1st place	2311 Gflop/s	IBM BG/L	DOE/NNSA/LLNL	Tom Spelce
1st runner up	1122 Gflop/s	Cray XT3 Dual	ORNL	Jeff Larkin
2nd runner up	1118 Gflop/s	Cray XT3	SNL	Courtenay Vaughan

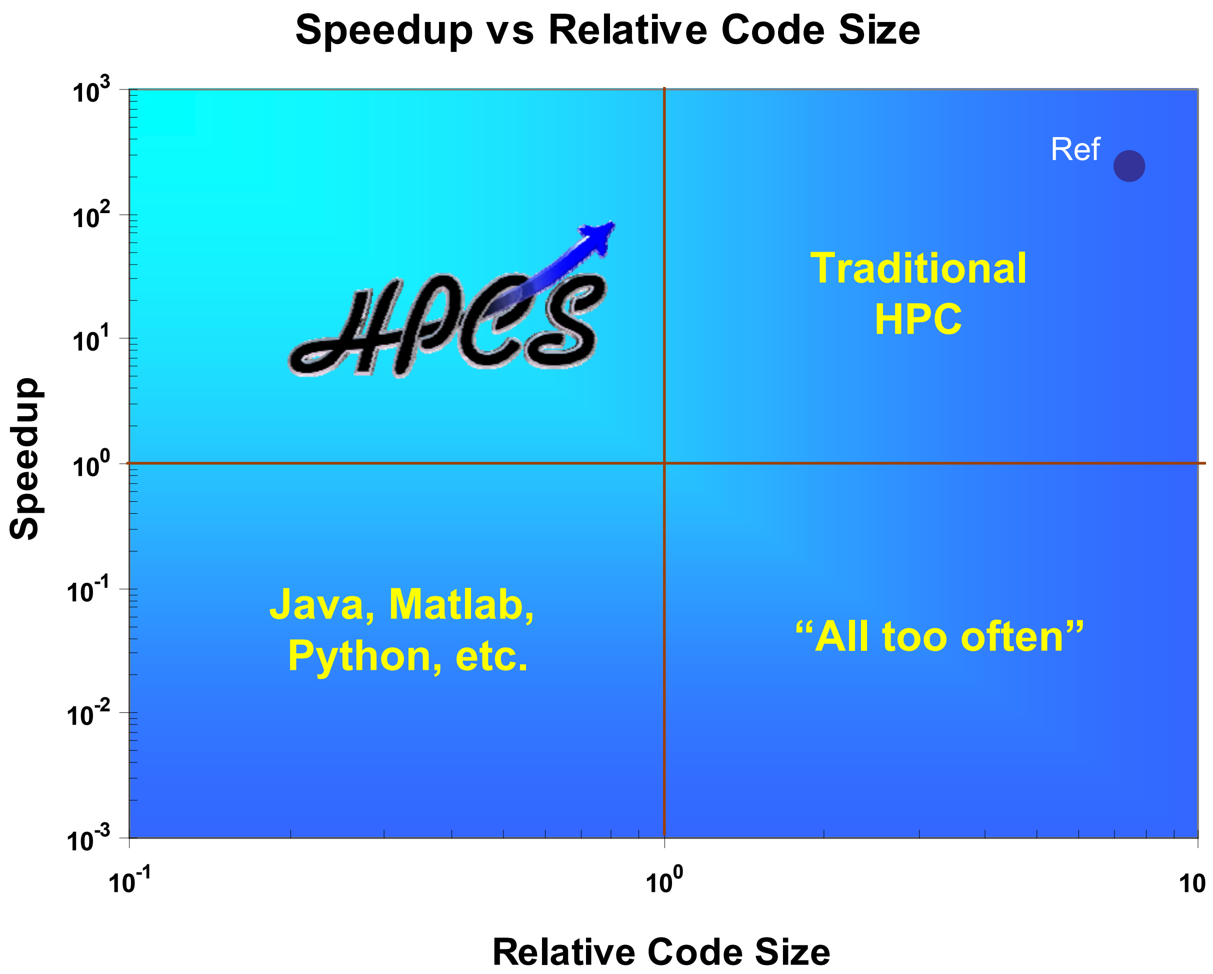
EP-STREAM-Triad (system)	Achieved	System	Affiliation	Submitter
1st place	160 TB/s	IBM BG/L	DOE/NNSA/LLNL	Tom Spelce
1st runner up	55 TB/s	IBM p5-575	LLNL	Charles Grassl
2nd runner up	43 TB/s	Cray XT3	SNL	Courtenay Vaughan

The Following are the Winners of the 2006 HPC Challenge Class 2 Awards

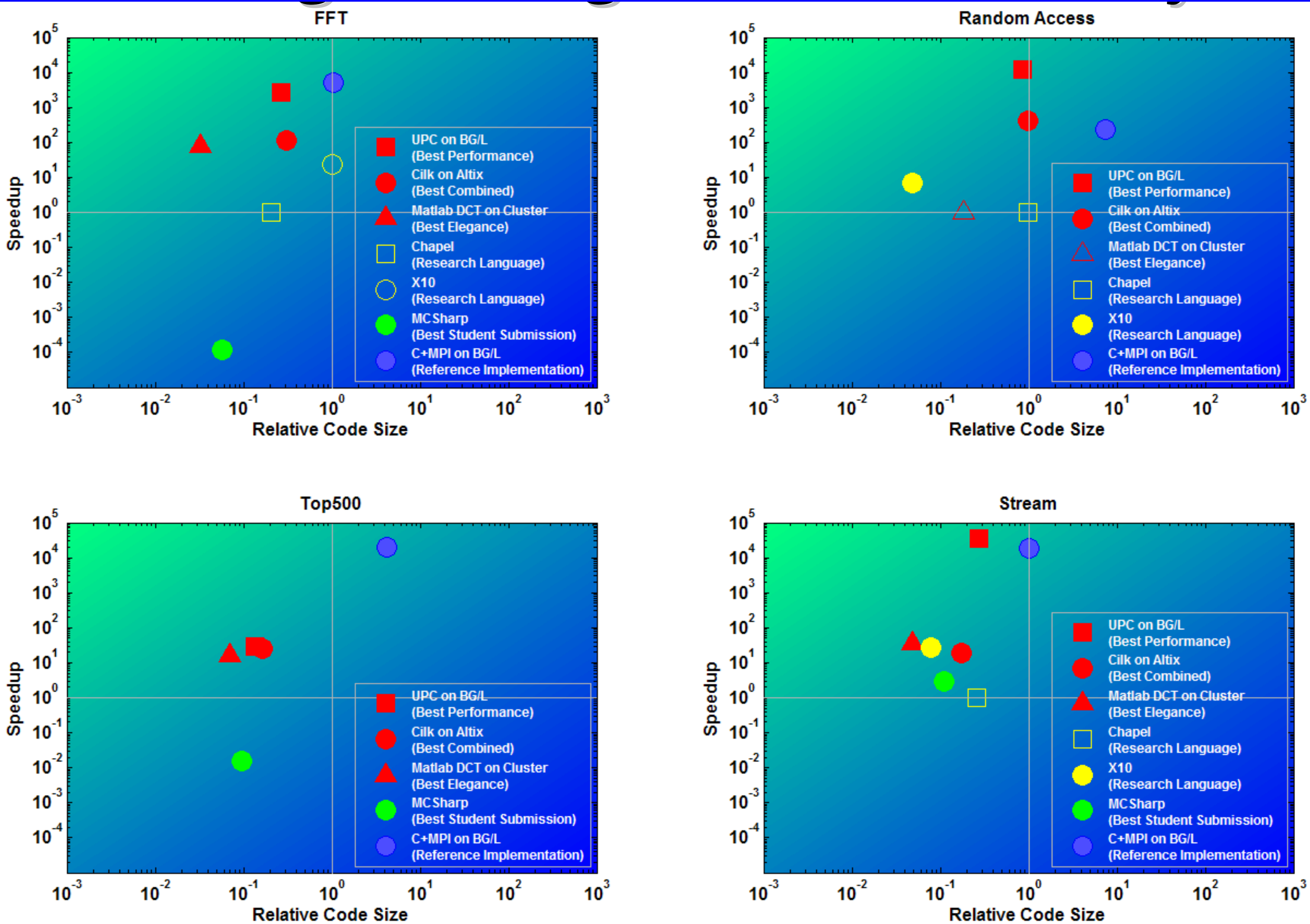
Award	Recipient	Affiliation	Language
Best Overall Productivity	Bradley Kuszmaul	MIT CSAIL	Cilk
Best Productivity in Performance	Calin Cascaval	IBM	UPC
Best Productivity and Elegance	Cleve Moler	MathWorks	Parallel MATLAB
Best Student Paper	Vadim Guzev	Russian People Friendship University	MC#
Honorable Mention	Brad Chamberlain	Cray Inc.	Chapel
Honorable Mention	Vivek Sarkar	IBM	X10

2006 Programmability

- Class 2 Award
  - 50% Performance
  - 50% Elegance
- 21 Codes submitted by 6 teams
- Speedup relative to serial C on workstation
- Code size relative to serial C



2006 Programming Results Summary



- 21 of 21 smaller than C+MPI Ref; 20 smaller than serial
- 15 of 21 faster than serial; 19 in HPCS quadrant

Top500 and HPC Challenge Rankings

- It should be clear that the HPL (Linpack Benchmark - Top500) is a relatively poor predictor of overall machine performance.
- For a given set of applications such as:
  - Calculations on unstructured grids
  - Effects of strong shock waves
  - Ab-initio quantum chemistry
  - Ocean general circulation model
  - CFD calculations w/multi-resolution grids
  - Weather forecasting
- There should be a different mix of components used to help predict the system performance.

Will the Top500 List Go Away?

- The Top500 continues to serve a valuable role in high performance computing.
  - Historical basis
  - Presents statistics on deployment
  - Projection on where things are going
  - Impartial view
  - Its simple to understand
  - Its fun
- The Top500 will continue to play a role

No Single Number for HPCC?

- Of course everyone wants a single number.
- With HPCC Benchmark you get 188 numbers per system run!
- Many have suggested weighting the seven tests in HPCC to come up with a single number.
  - LINPACK, MatMul, FFT, Stream, RandomAccess, Ptranspose, bandwidth & latency
- But your application is different than mine, so weights are dependent on the application.
- Score = W_1*LINPACK + W_2*MM + W_3*FFT+ W_4*Stream +

  W_5*RA + W_6*Ptrans + W_7*BW/Lat
- Problem is that the weights depend on your job mix.
- So it make sense to have a set of weights for each user or site.

Tools Needed to Help With Performance

- A tools that analyzed an application perhaps statically and/or dynamically.
- Output a set of weights for various sections of the application
  - [W_1, W_2, W_3, W_4, W_5, W_6, W_7, W_8]
  - The tool would also point to places where we were missing a benchmarking component for the mapping.
- Think of the benchmark components as a basis set for scientific applications
- A specific application has a set of "coefficients" of the basis set.

- Score = W_1*HPL + W_2*MM + W_3*FFT+ W_4*Stream + W_5*RA + W_6*Ptrans + W_7*BW/Lat + …

Future Directions

- Looking at reducing execution time
- Constructing a framework for benchmarks
- Developing machine signatures
- Plans are to expand the benchmark collection
  - Sparse matrix operations
  - I/O
  - Smith-Waterman (sequence alignment)
- Port to new systems
- Provide more implementations
  - Languages (Fortran, UPC, Co-Array)
  - Environments
  - Paradigms

Collaborators

- **HPC Challenge**
  - **Piotr Łuszczek, U of Tennessee**
  - **David Bailey, NERSC/LBL**
  - **Jeremy Kepner, MIT Lincoln Lab**
  - **David Koester, MITRE**
  - **Bob Lucas, ISI/USC**
  - **Rusty Lusk, ANL**
  - **John McCalpin, IBM, Austin**
  - **Rolf Rabenseifner, HLRS Stuttgart**
  - **Daisuke Takahashi, Tsukuba, Japan**
- **Top500**
  - **Hans Meuer, Prometeus**
  - **Erich Strohmaier, LBNL/NERSC**
  - **Horst Simon, LBNL/NERSC**

http://icl.cs.utk.edu/hpcc/