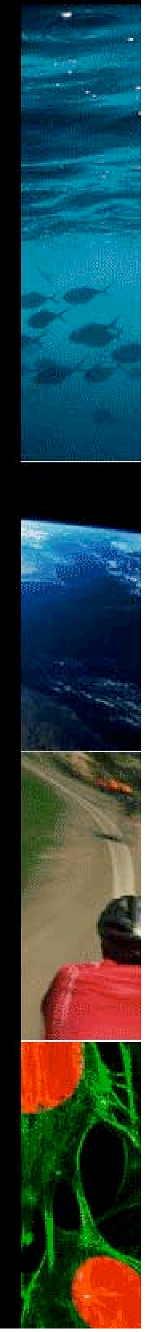
The Supercomputer Company

Cray and HPCC: Benchmark developments and results from the past year

Nathan Wichmann

wichmann@cray.com



CUG 2005

Outline

- What is HPCC
- What has changed in the past year?
- Results for Cray machines and competitive platforms
- Comparisons using HPCC
- Conclusions

HPCChallenge Project Goals

- To examine the performance of HPC architectures using kernels with more challenging memory access patterns than HPL (Linpack).
- To augment the Top500 list
- To provide benchmarks that bound the performance of many real applications as a function of memory access characteristics.

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

HPC Challenge Benchmark

- HPCC is as introduced (Nov’03) benchmark consisting of the following six main tests:
  - HPL the Linpack TPP benchmark which measures the floating point rate of execution for solving a linear system of equations.
  - PTRANS (parallel matrix transpose), exercises the communications where pairs of processors communicate with each other simultaneously.
  - STREAM, a simple synthetic benchmark program that measures sustainable memory bandwidth (in GB/s) and the corresponding computation rate for simple vector kernel.
  - RandomAccess, measures the rate of random integer updates of memory.
  - ^{b}eff (MPI bandwidth & latency test), a set of tests to measure the latency and bandwidth of a number of simultaneous communication patterns.

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

HPCC Benchmark: What changed?

- Last year HPCC added the following test
  - DGEMM measures the floating point rate of execution of double precision real matrix-matrix multiplication
  - FFTE measures the floating point rate of execution of double precision complex one-dimensional Fast Fourier Transform .
  - Global Random Access changed the base algorithm to not “sort” but to use mpi_isends and mpi_irecvs from “ANY_SOURCE”
    - Dramatic negative effect on base performance
    - Original version was considered “not in the spirit of the benchmark”

Cray: Introduced 3 Machines

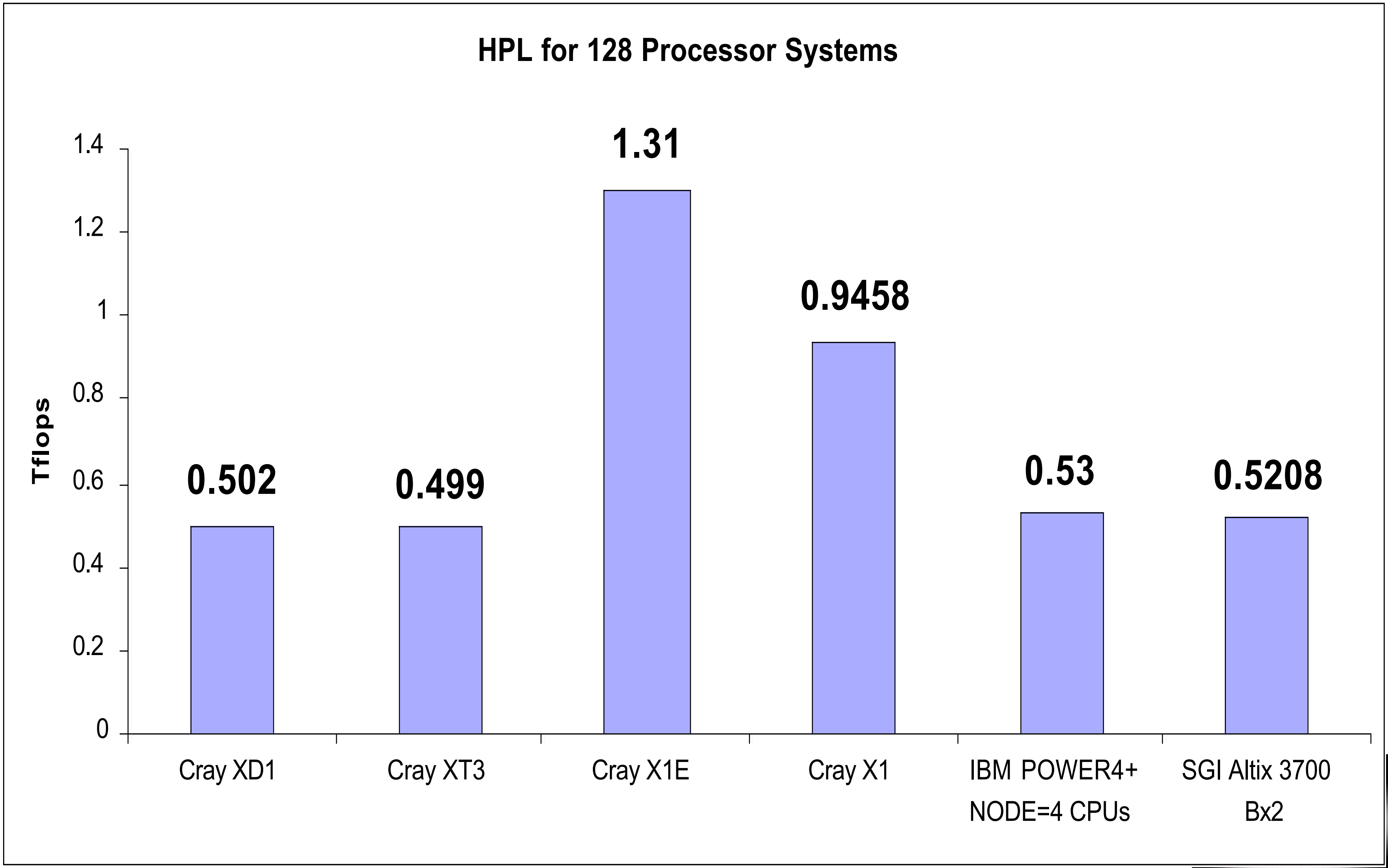
- Cray XD1
  - Supercomputer performance at a reasonable price
  - 2 socket Opteron SMP
  - Linux OS
- Cray XT3
  - High Scalability MPP
  - Single Opteron per node
  - Catamount OS
- Cray X1E
  - Vector Processor follow on to the X1
  - Triple the Peak Price/Performance
  - Improved scalar performance

Results for similar number of CPUs

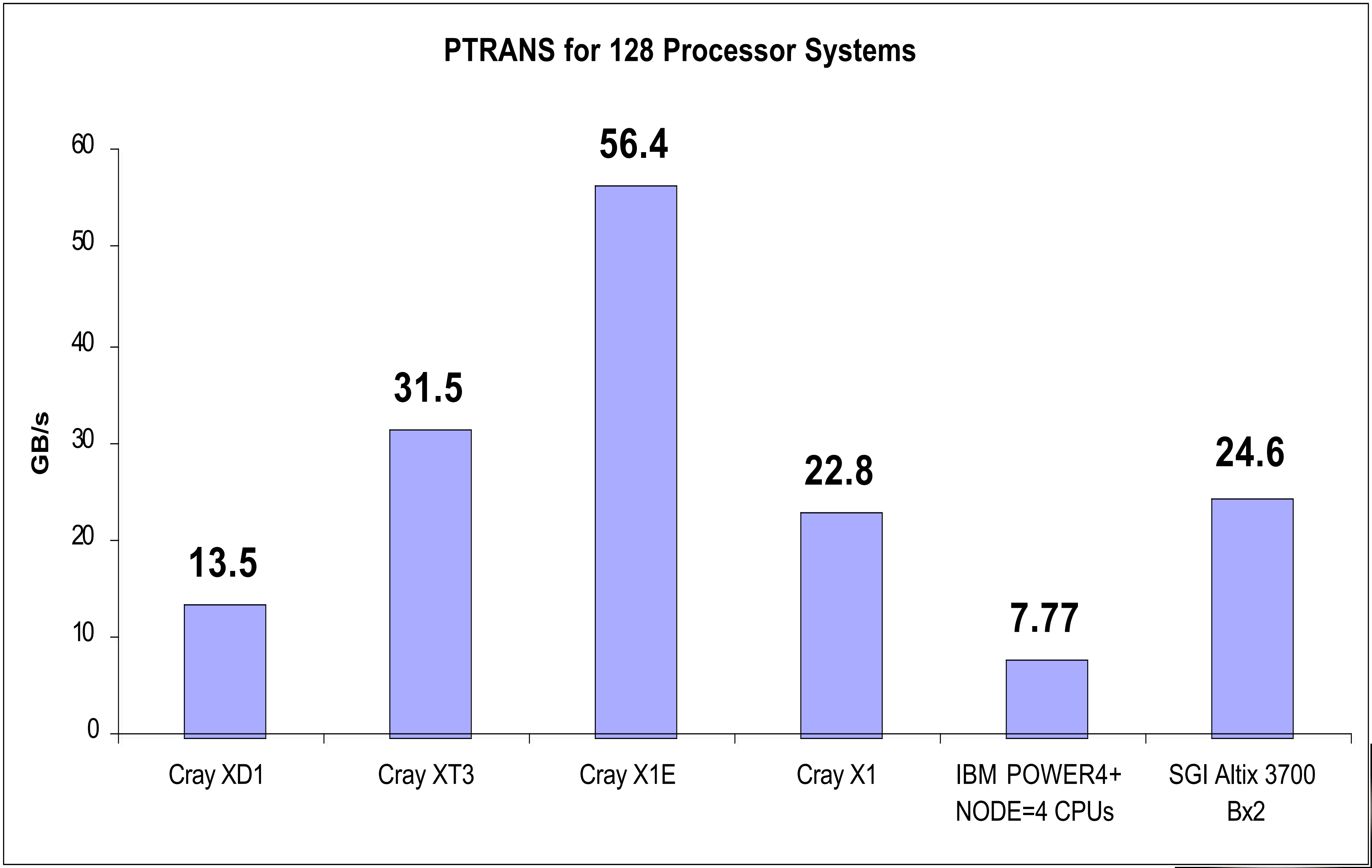
- 6 machines, all with approx 128 cpus

Machine	CPU type	Ghz	NCPUs
Cray XD1	Opteron	2.4 Ghz	128
Cray XT3	Opteron	2.4 Ghz	128
Cray X1E	Vector MSP	1.13 Ghz	124
Cray X1	Vector MSP	.8 Ghz	124
IBM	POWER4+	1.7 Ghz	128
SGI Altix 3700 Bx2	Itanium2	1.6 Ghz	128

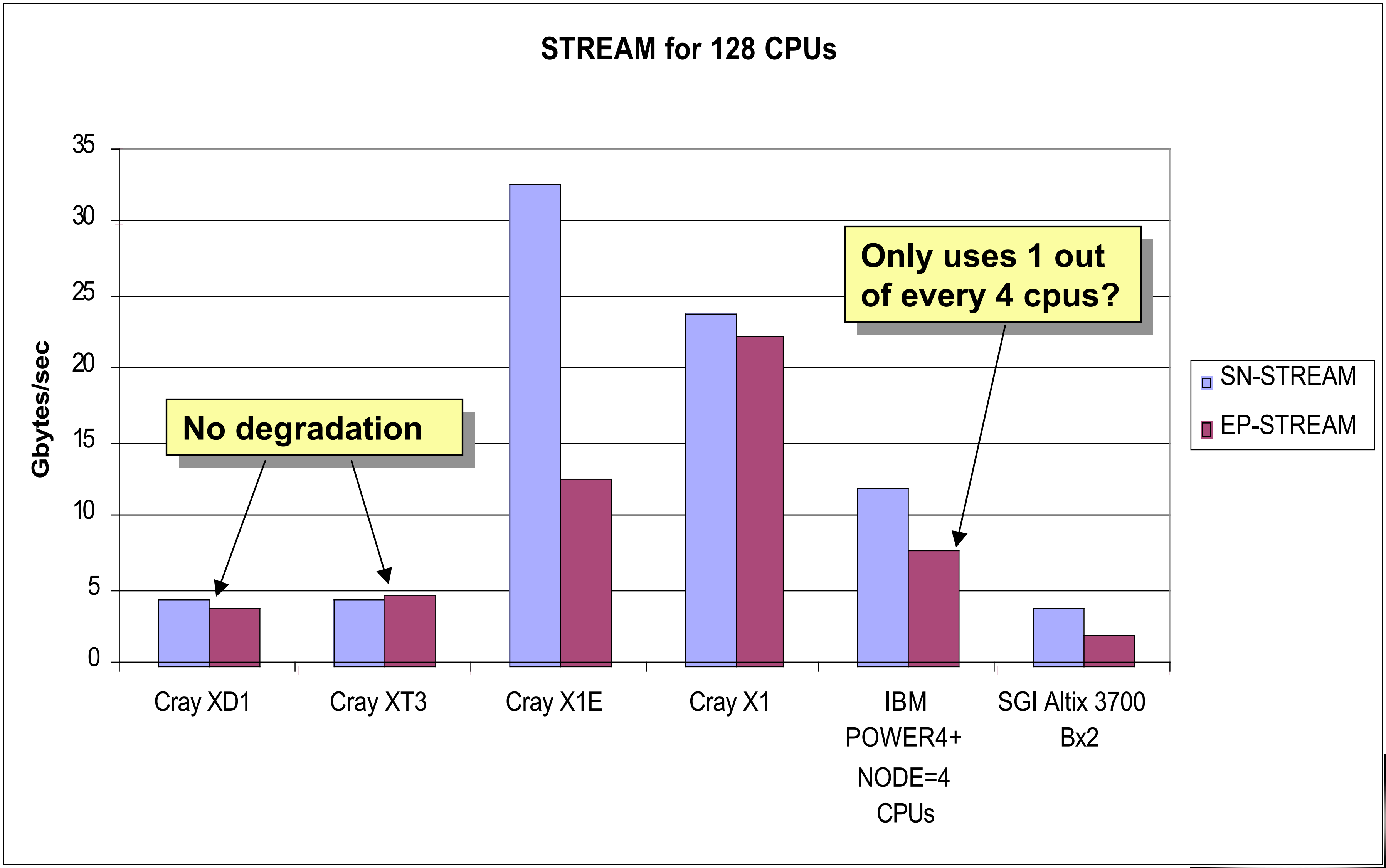
HPL Results


HPL for 128 Processor Systems
Tflops
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0.502
0.499
1.31
0.9458
0.53
0.5208
Cray XD1
Cray XT3
Cray X1E
Cray X1
IBM POWER4+ NODE=4 CPUs
SGI Altix 3700 Bx2

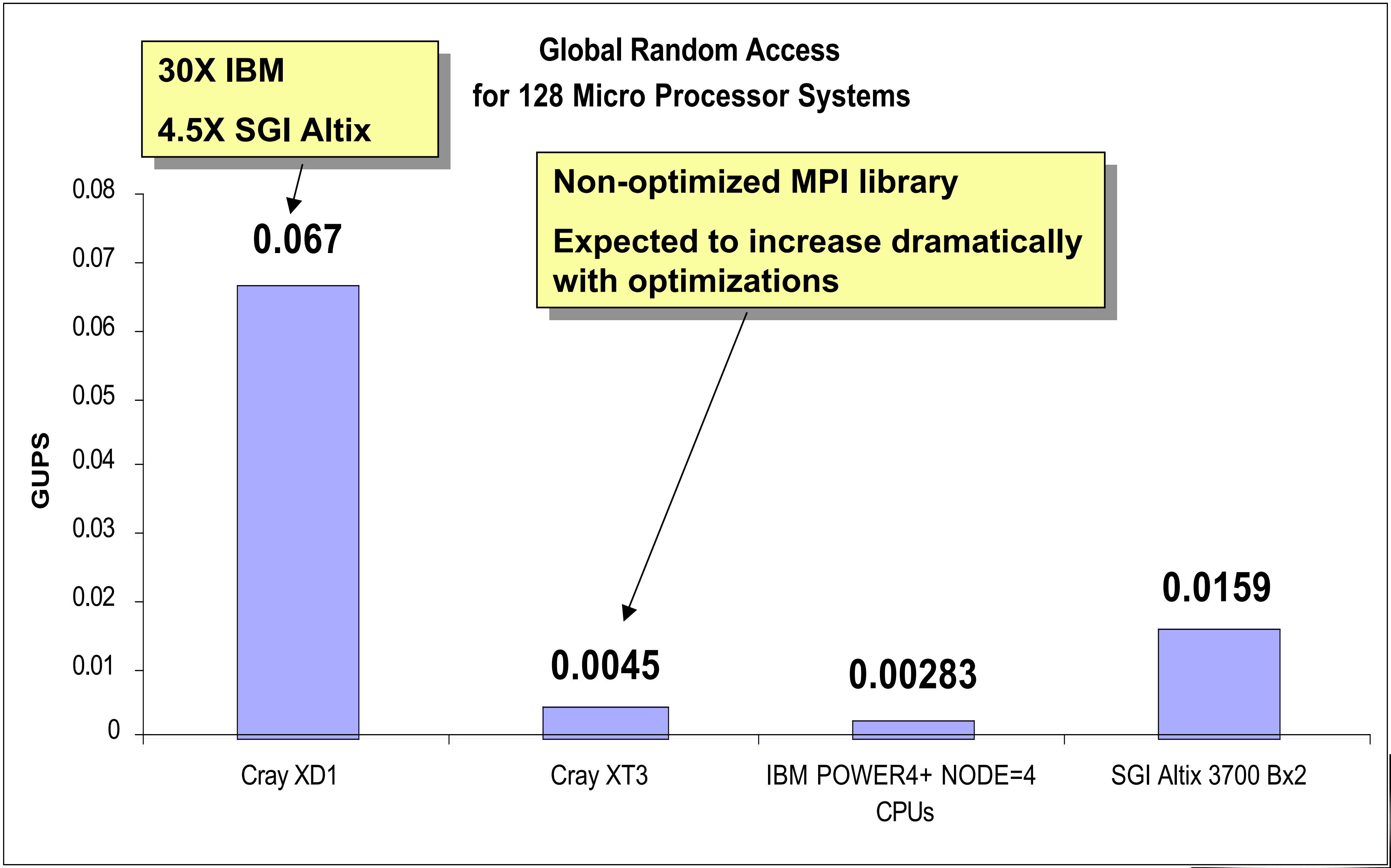
PTRANS Results



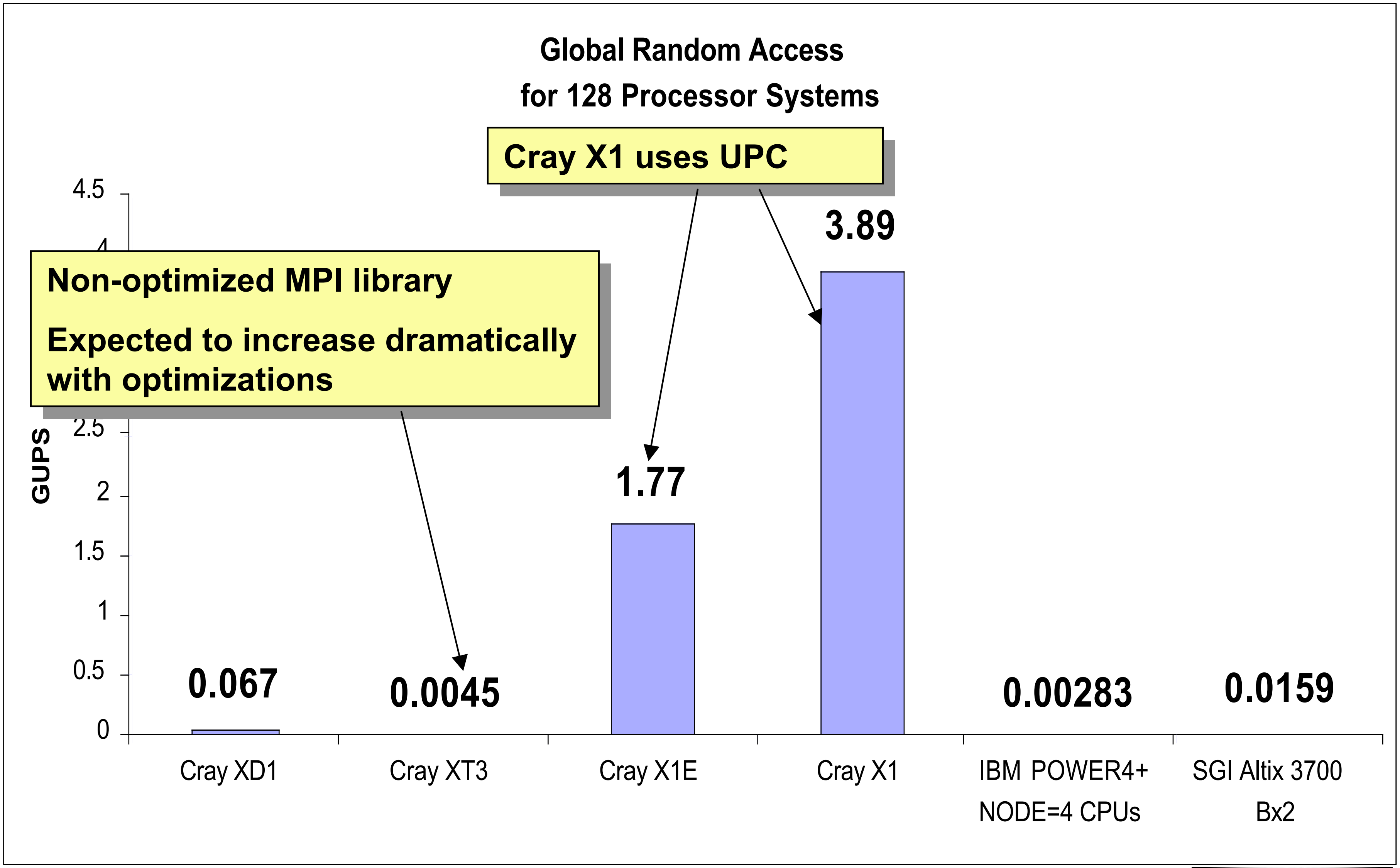
STREAM Results



Global Random Access Results



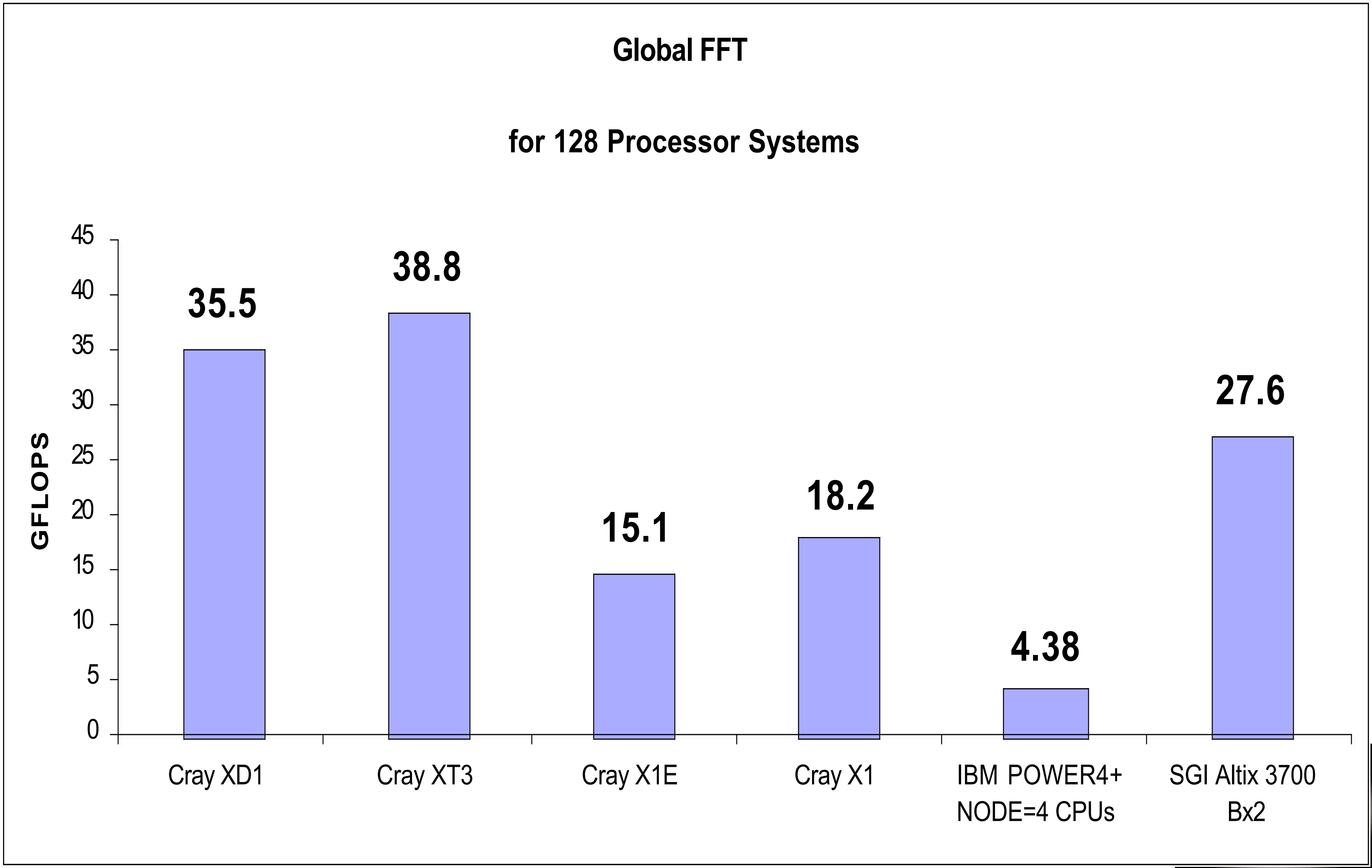
Global Random Access Results



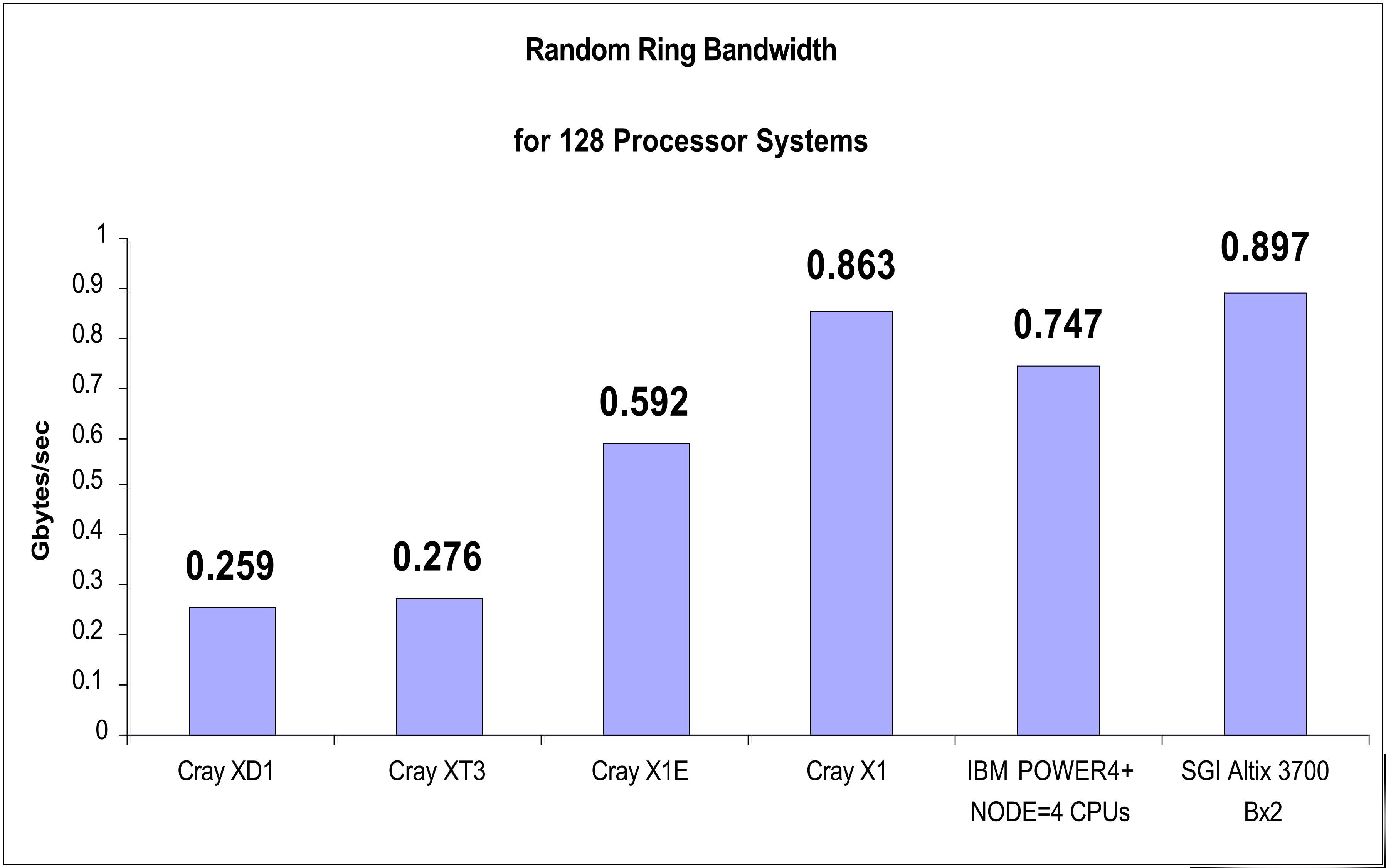
Problems with G-FFT

- We have seen problems with the G-FFT test
- Error occurs on at least 3 platforms
  - XD1, XT3, X1(E)
  - Not sure it is the same problem but looks that way
- Seems to run small problems ok, but has more difficulty as problem size and/or number of CPUs grow
- Seems to be an error in the FFTW code
  - Downloaded FFTW source and have similar problems
  - Fails in initialization
- Working internally and with HPCC authors to debug and correct the situation

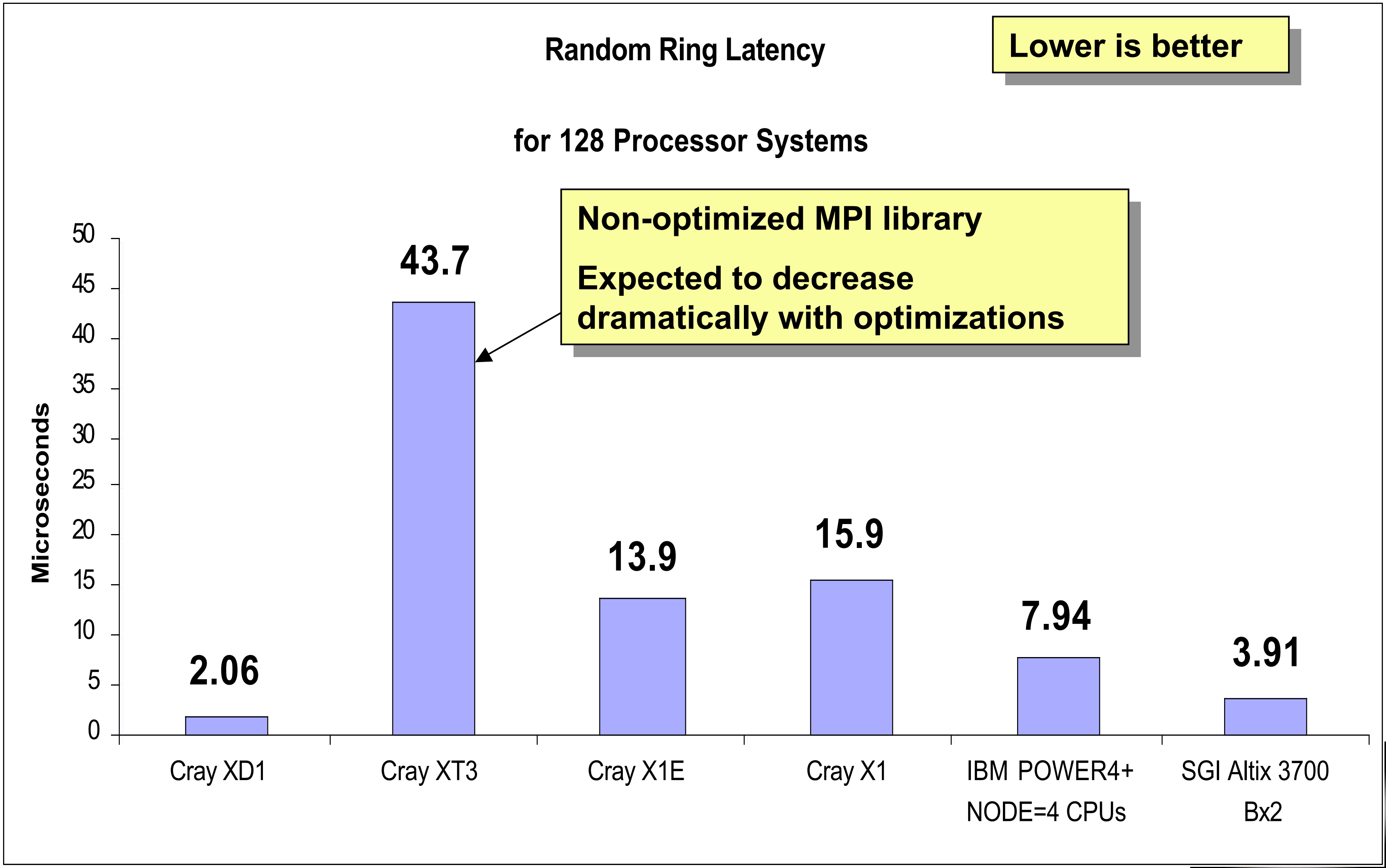
Global FFT Results



Random Ring Bandwidth Results



Random Ring Latency Results

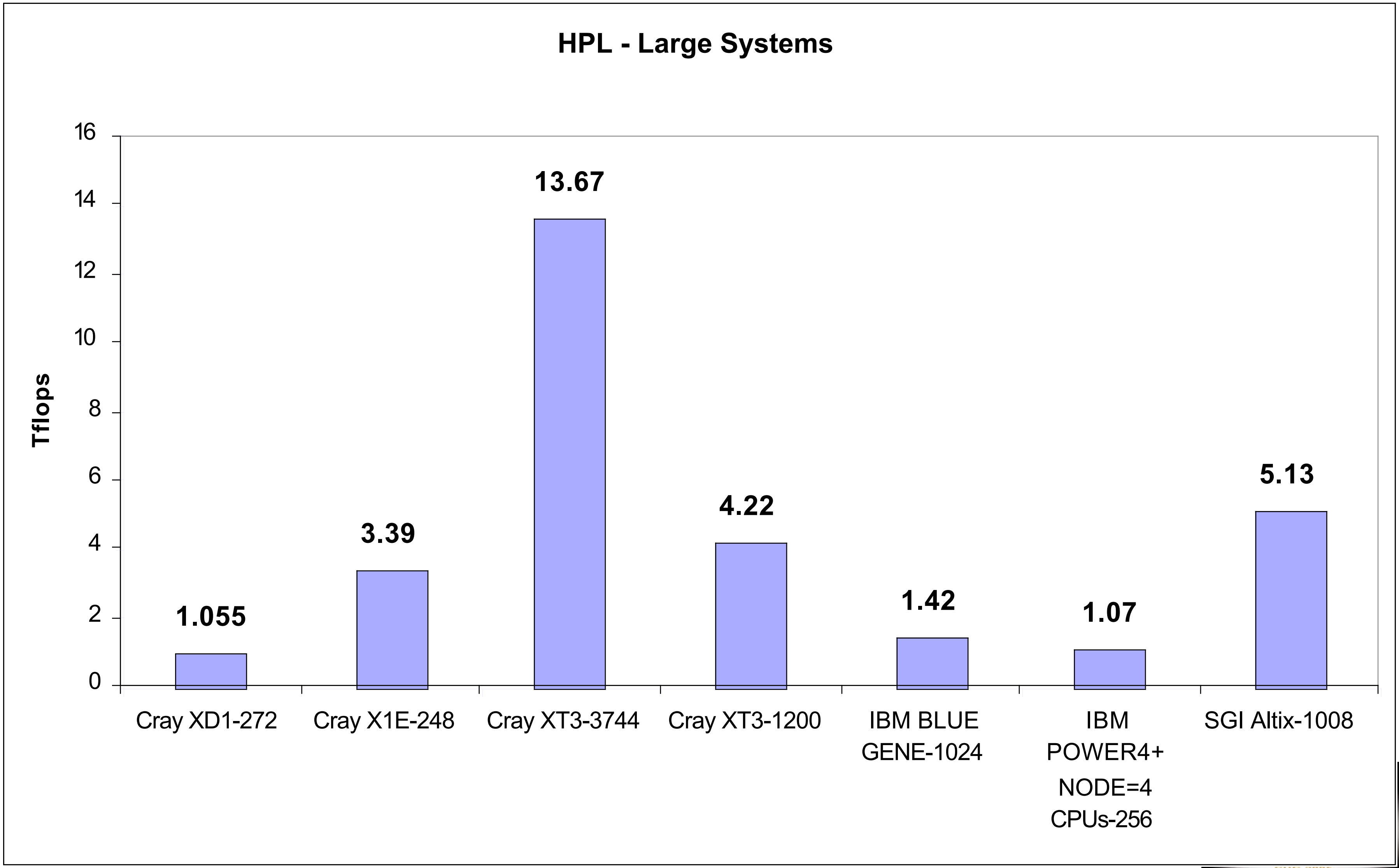


Results for similar number of CPUs

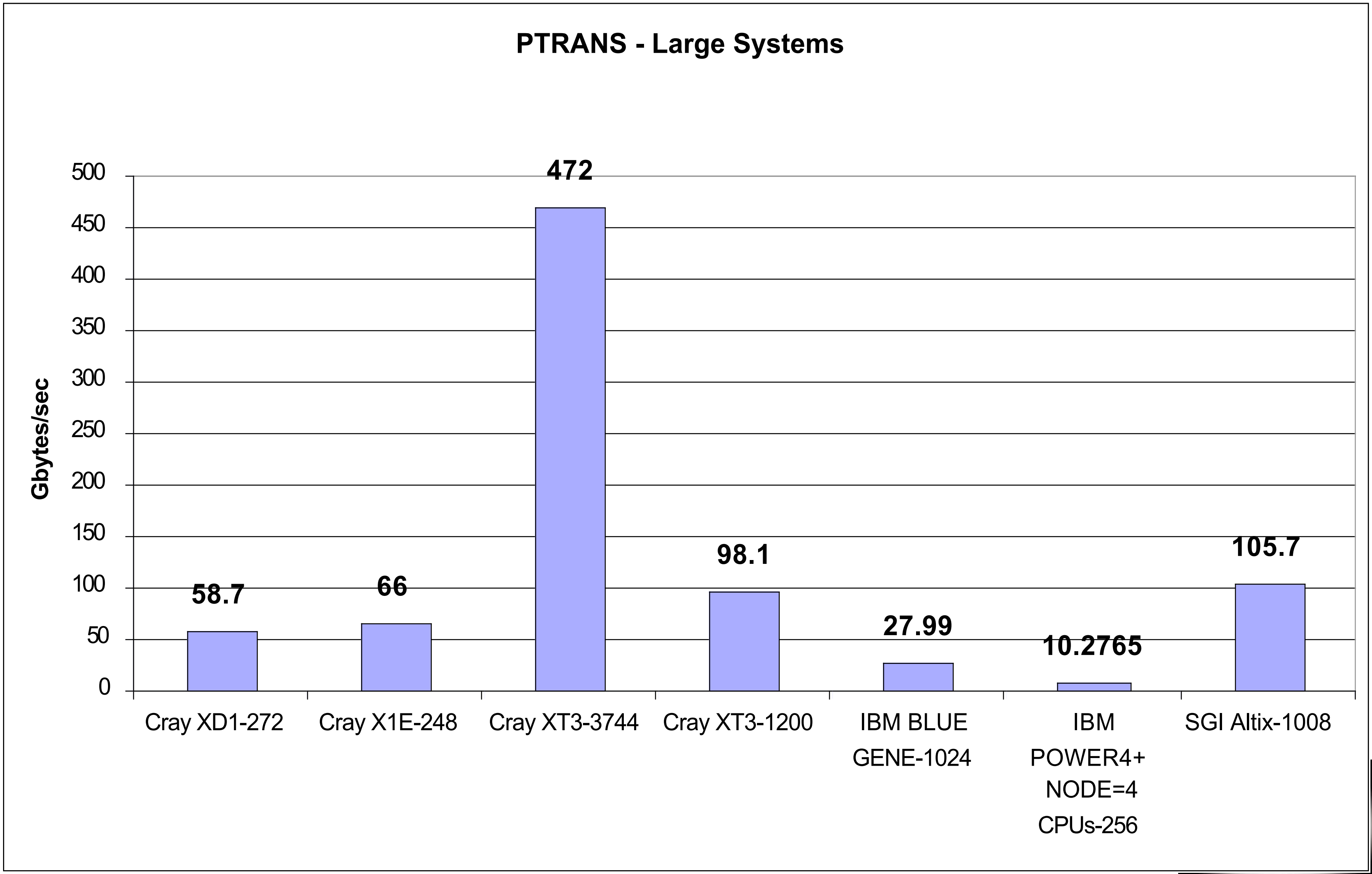
- 7 machines, using the largest configurations I could find

Machine	CPU type	Ghz	NCPUs
Cray XD1	Opteron	2.4 Ghz	272
Cray XT3	Opteron	2.4 Ghz	3744 & 1200
Cray X1E	Vector MSP	1.13 Ghz	248
IBM Blue Gene	PowerPC	.44 Ghz	1024
IBM	POWER4+	1.7 Ghz	256
SGI Altix 3700 Bx2	Itanium2	1.6 Ghz	1008

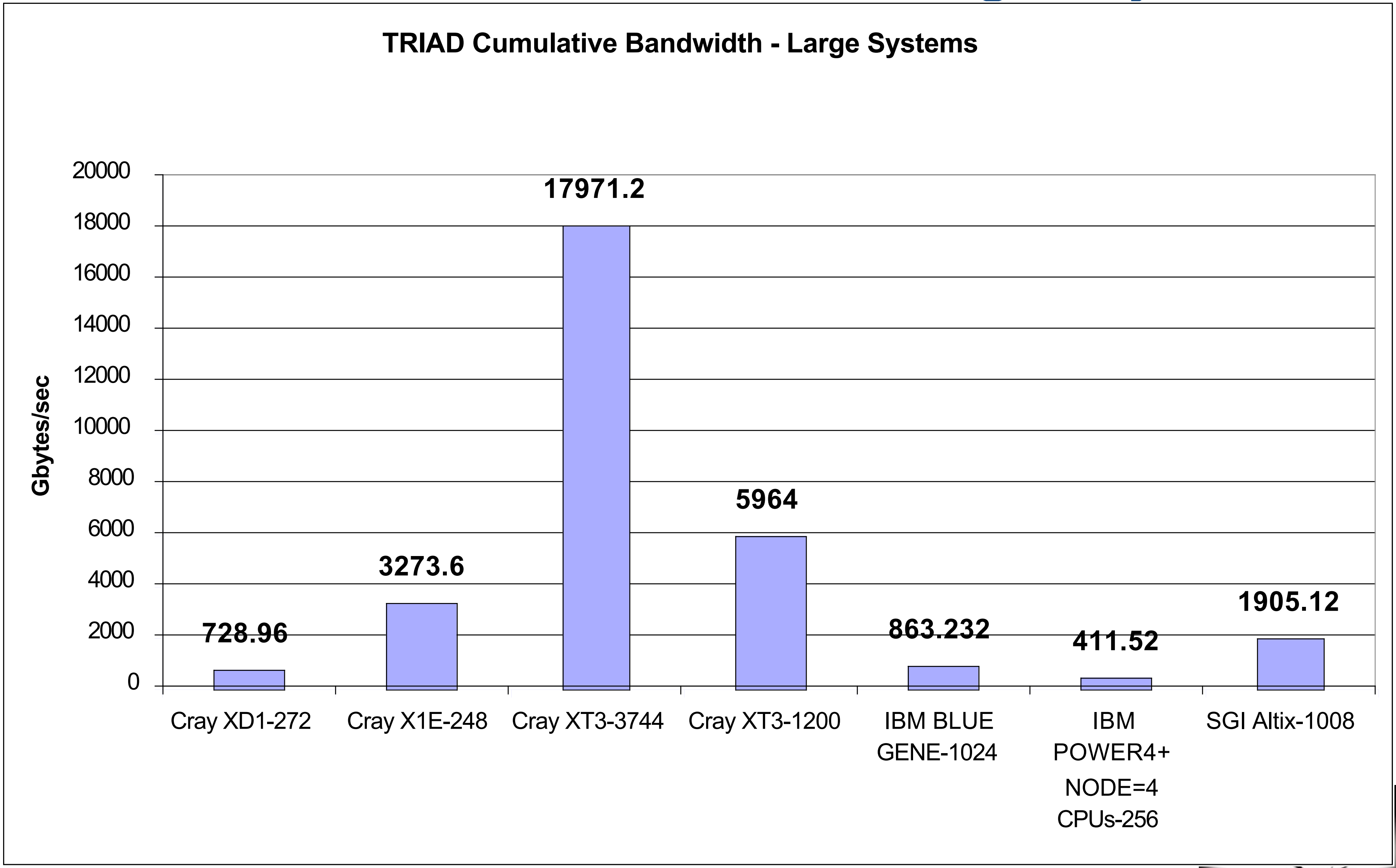
HPL Results for Large Systems



PTRANS Results for Large Systems



Cumulative Bandwidth for Large Systems

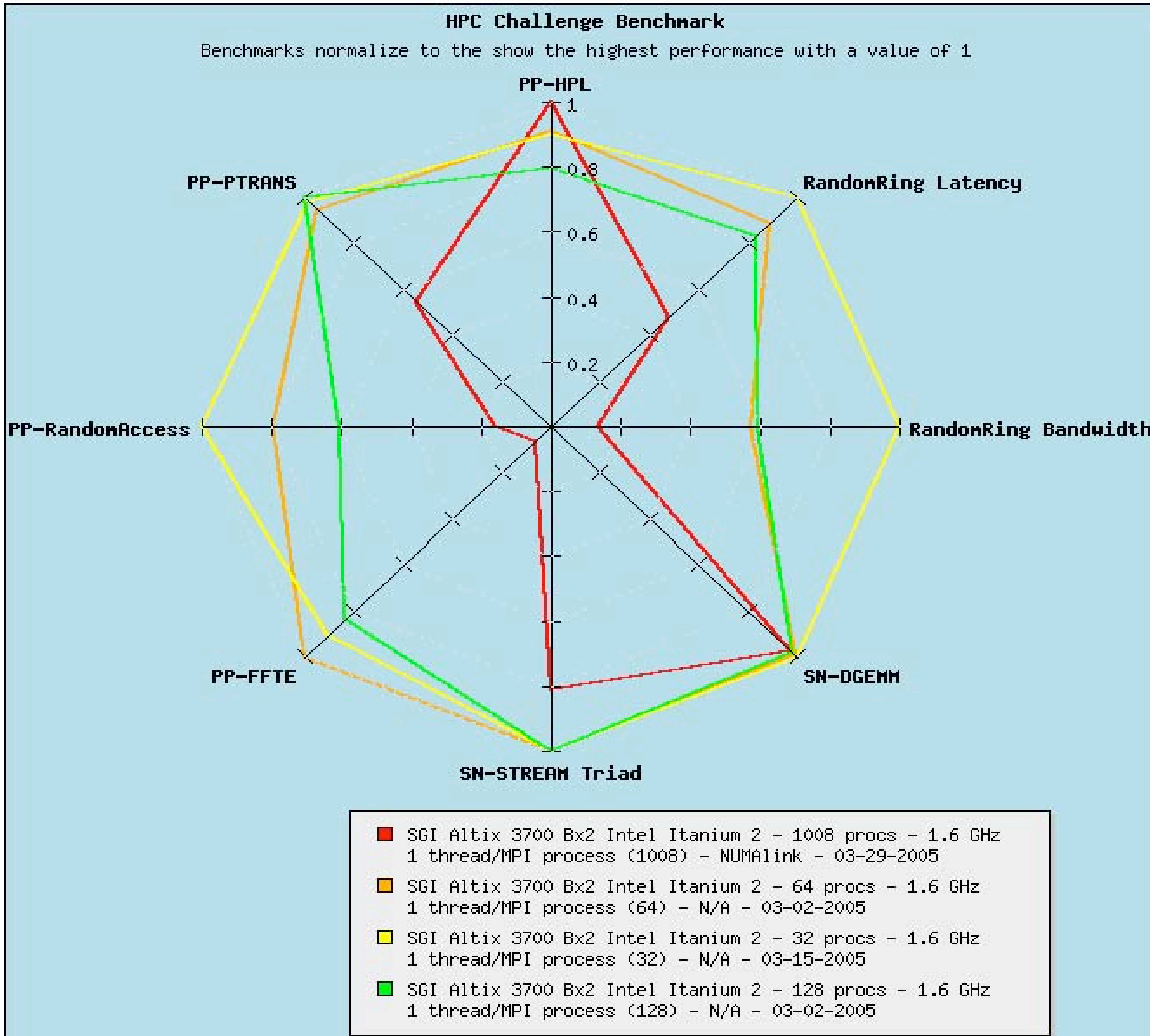


Comparing Machines Using Kiviat Diagrams

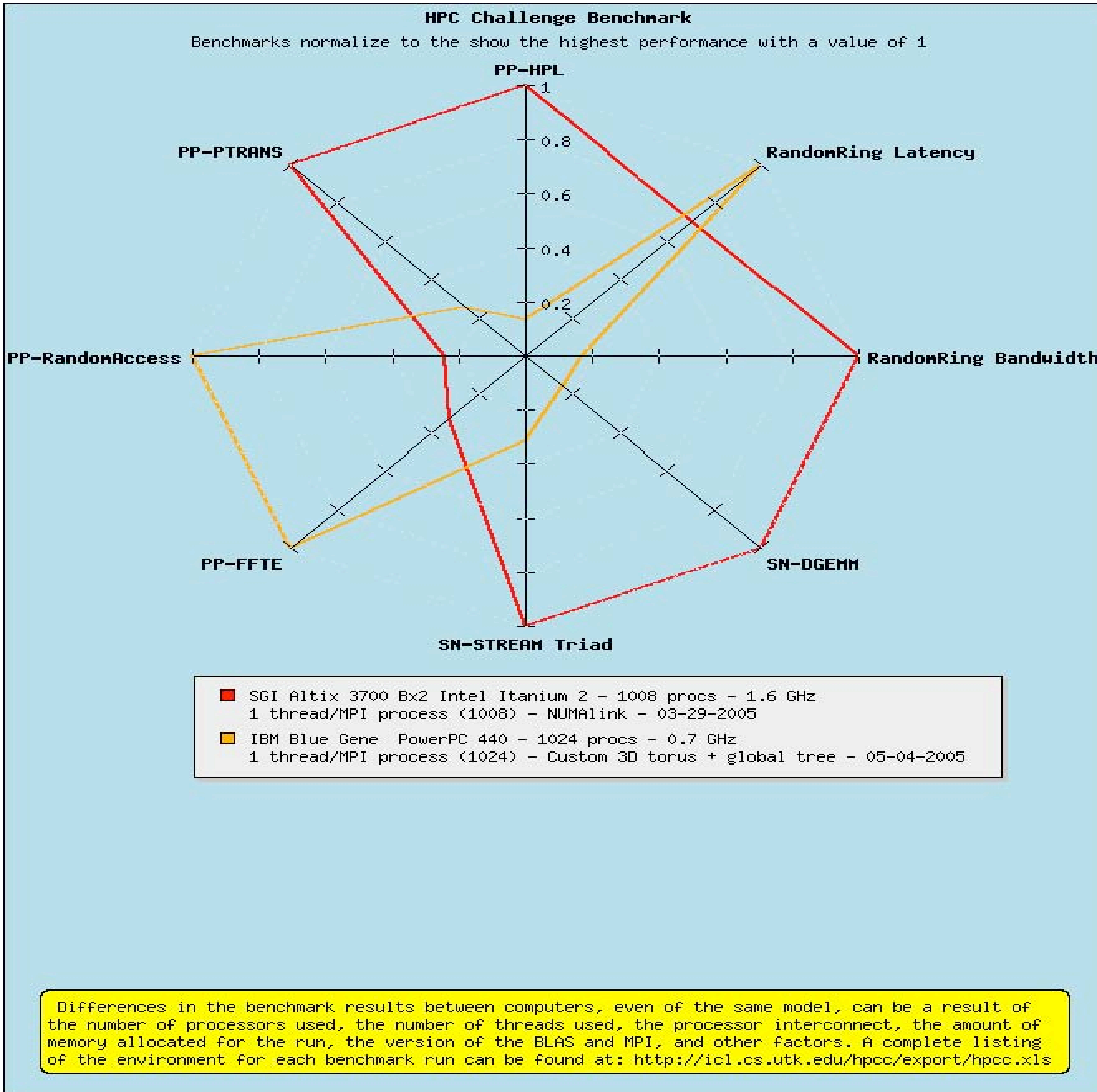
- HPCC web site now allows each machine to be plotted on a Kiviat Diagram
- A Kiviat Diagram is a 2 dimensional graph of radially extending lines where each line corresponds to a test
  - To create a diagram, each score first is turned into a Per Processor metric
  - Scores are then normalized to the best score in each category
  - A perimeter is drawn connecting all the point for a given machine

Kiviat Diagrams (cont.)

- Pros
  - Visually simple
    - Strongly implies the larger the area inside the perimeter, the "better" the machine
- Cons
  - Per Processor results end up making smaller machine look better
    - Cannot use diagram to determine which is the more "powerful" machine
  - Are all tests created equal?
    - How do you graph RR Latency on the same graph as HPL?
  - Does not allow one to plot optimized results
  - Left to "eyeball" the results



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


HPC Challenge Benchmark
Benchmarks normalize to the show the highest performance with a value of 1
PP-HPL
1
0.8
0.6
0.4
0.2
RandomRing Latency
RandomRing Bandwidth
SN-DGEMM
SN-STREAM Triad
PP-FFTE
PP-RandomAccess
PP-PTRANS
SGI Altix 3700 Bx2 Intel Itanium 2 - 1008 procs - 1.6 GHz
1 thread/MPI process (1008) - NUMAlink - 03-29-2005
IBM Blue Gene PowerPC 440 - 1024 procs - 0.7 GHz
1 thread/MPI process (1024) - Custom 3D torus + global tree - 05-04-2005
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

Comparing Machines Using the App Kernel Power Rating

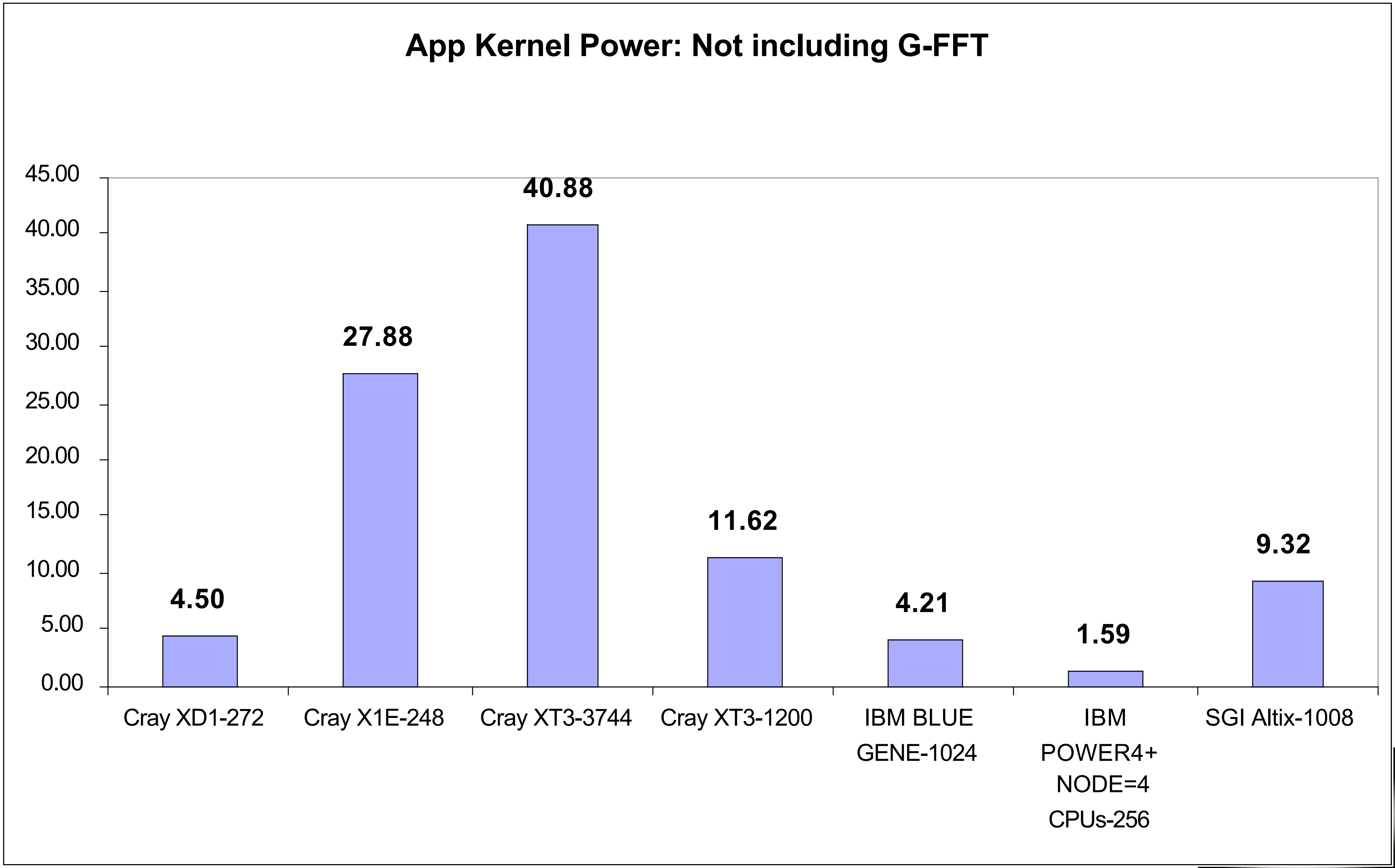
- Cray has developed a internal rating to help in interpreting HPCC results which I call the App Kernel Power Rating
- Normalize scores
  - In each category take test result and divide by the combined power of all machines
    - Creates a unitless number
    - Equal to a percentage of total power
- Combine the 5 application kernels into 1 number using equal weighting
  - HPL; PTRANS; Cumulative BW; G-RA, and G-FFT
  - Every kernel actually does useful work
  - Excellent cross section of real applications
  - Easy to combine into one number

App Kernel Power Rating (cont.)

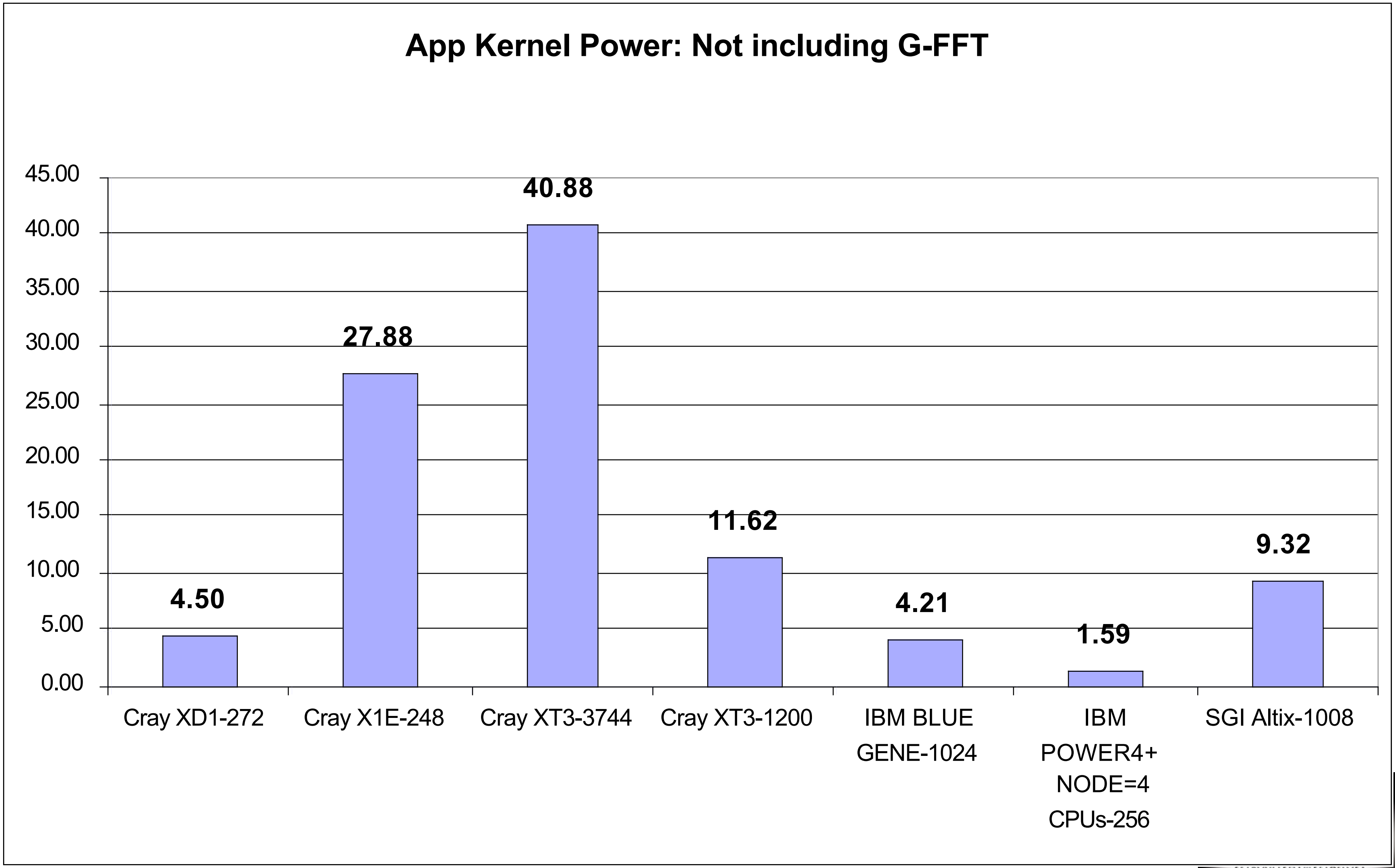
- Pros
  - Single number score
    - Obvious winner
  - Simple to understand
  - Focuses on kernels that do “real work”
  - Larger machines are considered more powerful
- Cons
  - Are all tests created equal?
    - RR Latency and Bandwidth are not included

- **VERY** similar to area calculation of a Kiviat Diagram using normalized but absolute score (not per processor)

App Kernel Power Ratings



App Kernel Power Ratings for Large Systems



Conclusions

- When comparing similar number of cpus
  - When using the App Kernel Power Rating, the Cray XT3 and XD1 are 10-180% more powerful than a SGI or IBM POWER4+ system
  - When comparing SN-STREAM to EP-STREAM numbers, the XD1 and XT3 hold up better than any other platform
  - The Cray XD1 has exceptional Latency characteristics.
  - The exceptional Global Random Access performance of the X1(E) makes its power rating look very good.

Conclusions (cont.)

- Machine size and pricing vary, so it is difficult to make an apples-to-apples comparison; but let's try
  - HPL: Blue Gene efficiency and peak are very low
  - PTRANS:
    - The XT3 holds the "world record" of 472 Gbytes/sec
    - The XD1 is very high for 272 processor
    - IBM systems trail substantially
  - TRIAD Cumulative Bandwidth:
    - The Cray XT3's combination of excellent per processor bandwidth and large cpu counts make it stand out
  - App Kernel Power Rating
    - The X1E's exceptional GUPS score keeps it towards the top
    - The XT3's consistent power gives it a very high rating
    - The a 272 processor XD1 is more powerful than 1024 IBM Blue Gene
- HPCC is a powerful new tool for examining machine performance using more challenging kernels

PETROGLYPHS to PETAFLOPS
CUG 2005