Performance Predictions of NAS Parallel Benchmarks on HPC Platforms

using MAPS and MetaSim

developed by the Performance Modeling and Characterization (PMaC) Laboratory

MISSION: **to bring scientific rigor to the prediction of scientific application performance on current and projected HPC platforms.**

OVERVIEW: This research currently focuses on developing tools and techniques to predict the performance of the compute intensive kernels of the NAS Parallel Benchmarks (NPBs) on current HPC platforms (SDSC Blue Horizon, PSC Compaq TCSini, UT Cray T3E, and SDSC IA-64 Itanium cluster). Future research will utilize these tools and techniques for performance prediction of scientific applications.

Our technique is to measure each factor that affects performance of an application in isolation from any particular machine; we call this our ***Application Signature***. This is a detailed summary of the fundamental operations to be carried out by the application, independent of a machine.

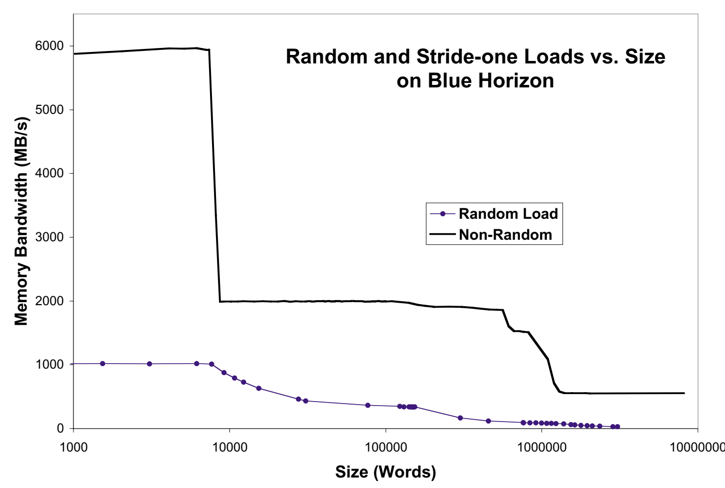
Next we develop a ***Machine Signature***, the characterization of the rates at which a machine can carry out fundamental operations independent of a particular application.

The ***Convolution Method*** is the algebraic mapping of the Application Signature on to the Machine Signature to arrive at a performance prediction

In this research we begin with simple models and add complexity as needed. Thus, the performance prediction of the kernels is guided by two "rules of thumb":

1. The ***per-processor and/or single processor performance*** of an application is primarily a function of how it interacts with the local memory hierarchy.
2. The ***scalability*** of an application is primarily a function of the way it interacts with the interconnect.

Guided by the first rule, the single processor Machine Signatures and Application Signatures are based on a ***local memory performance*** and ***use***.



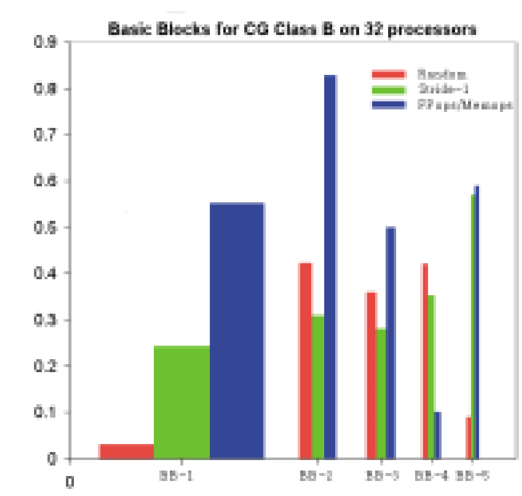
MACHINE SIGNATURE PRODUCED BY MAPS

This Machine Signature for SDSC's ***IBM Blue Horizon*** is based on the local memory performance measured using the ***Memory Access Patterns (MAPs)*** benchmark, designed to measure the bandwidth for loads and stores from all levels of memory. These measurements include stride one and random access. MAPs is hand tuned for each architecture.

CONVOLUTION METHOD

The Machine Signature and the Application Signature can be combined to produce the performance of a single processor, by using the simulated cache hit rates of an application (MetaSim) and combining them with the measured bandwidths for those hit rates (MAPs).

MAPs and MetaSim are used to calculate the per-processor performance; ***DIMEMAS***, a network simulator, predicts the performance of a parallel application. Dimemas uses ***MPI trace files*** collected from the application. Dimemas uses these trace files, network specifications of the new machine, and a ratio of the relative speeds of the new machine's processor to the processor on which the trace files were collected, to predict the performance of the application on the new machine.



APPLICATION SIGNATURE PRODUCED BY METASIM

This Application Signature for the ***NPB kernel CG*** was measured using ***MetaSim***, a semi-cycle accurate simulator capable of measuring an applications use of the memory hierarchy. Given a user-supplied description of a machine's memory subsystem, MetaSim can produce information about how an application uses the memory of such a system.

AN EXAMPLE: METASIM DATA FROM BLUE HORIZON

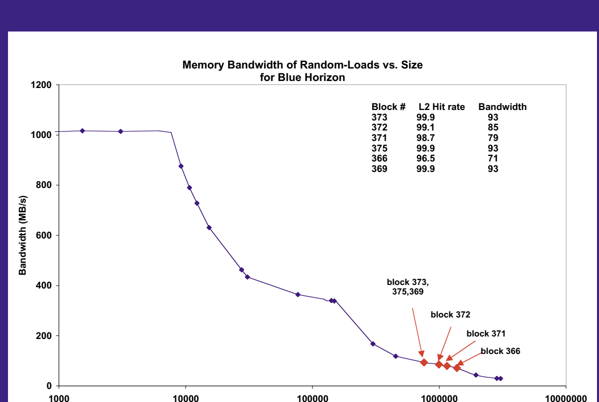
BLOCK NUMBER	NUMBER INSTRUCTIONS	MEMORY REFERENCES	% MEMORY REFERENCES	L1 HIT RATIO	L2 HIT RATIO	MAPS BANDWIDTH	RANDOM RATIO
373	2.05E+09	8.8E+08	0.2197	92.19	99.99	93	0.33
372	1.63E+09	8.3E+08	0.2067	90.17	99.14	85	0.37
371	1.15E+09	5.6E+08	0.1381	88.79	98.68	79	0.37
375	1.37E+09	5E+08	0.1242	93.02	99.99	93	0.35
366	4.99E+09	1.8E+08	0.0443	85.13	96.51	71	0.15
386	2.60E+09	1.7E+08	0.0431	96.15	99.34	1513	0
369	2.97E+09	1.5E+08	0.0369	93.43	99.99	93	0.51

Reference patterns for each basic block

Used to weight each basic block's bandwidth number

If >.1, use data from random-loads

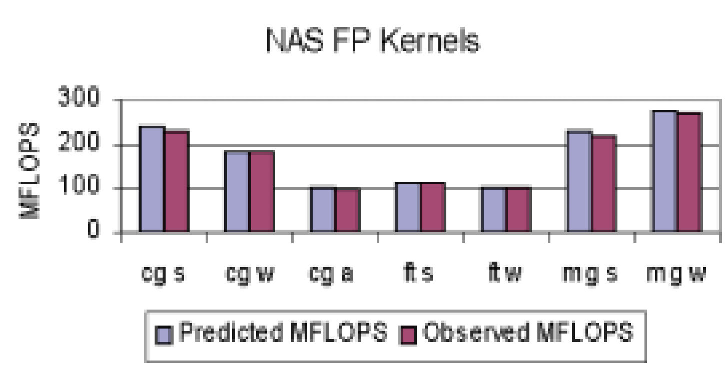
Used to determine bandwidth numbers from MAPs data



MetaSim generates the data about an application's use of the memory hierarchy for each basic block, which produces the application's relative performance on that machine. The data is then combined with the MAPS curves. The equation (at lower left) is used to sum the weighted bandwidths of each basic block to calculate its MetaSim number. This number is used as a processor speed ratio supplied to the DIMEMAS network simulator.

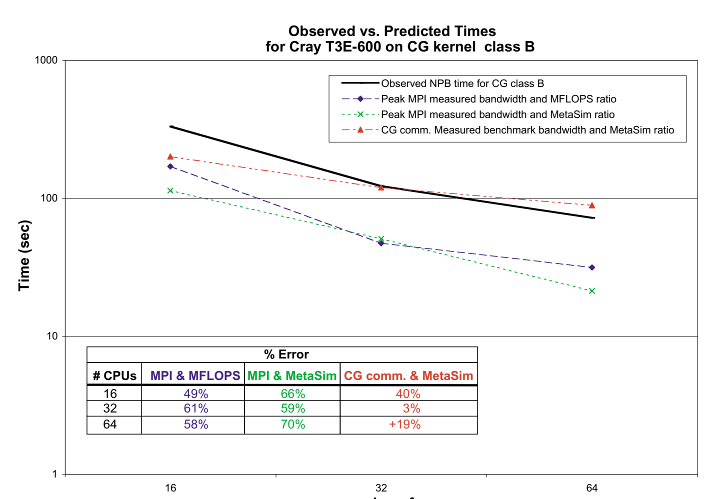
P=Processor speed based on memory bandwidth.

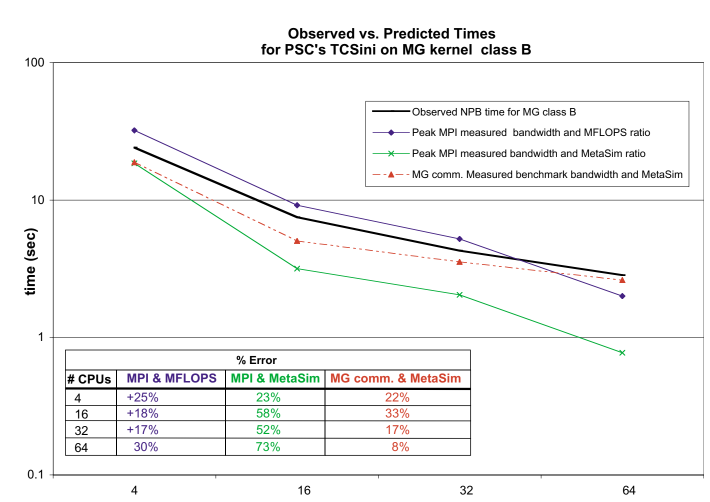
$$P = \sum_{i=\text{basic block}} (\text{\% of Total Memory references})_i \cdot (\text{MAPs Bandwidth})_i$$



PREDICTED PERFORMANCE OF SERIAL NPBs

The graph at left shows the prediction performance using MetaSim of the floating-point intensive kernels of the serial NPBs. These results show that for single processor performance, MetaSim does a good job of predicting the performance.





PREDICTED PERFORMANCE OF PARALLEL NPBs

CG and MG communication benchmarks were developed to adjust MPI peak bandwidth. These results illustrate that the use of MetaSim and the communication benchmarks predict better than peak numbers.

www.sdsc.edu/PMaC/

Sponsored by DOE Office of Science