

x86_64 Dual-core CPU を用いた PC cluster の性能評価

幸谷智紀*

大規模な科学技術計算を実行するための計算機環境は、さまざまな所で求められるようになってきている。特に大学研究室レベルにおいては、教育用にも使用できる高性能かつ汎用的なコンシューマ向け PC を組み合わせたコモディティ PC cluster が多く利用されている。2005 年には現行の x86_32 CPU と互換性を持ち、なおかつ 64bit 環境を提供できる Dual-Core CPU が AMD 及び Intel から登場し、コモディティ PC cluster でもより大規模な並列計算が実行できるようになりつつある。しかし、IEEE754 倍精度・多倍長数値計算及びその並列化に関する、これらの CPU 性能を調査した文献はまだ少ない。本稿では、コンシューマ向け PC 向けの x86_64 Dual-core CPU である Intel Pentium D 820 と AMD Athlon64 X2 3800+ の数値計算性能及びネットワーク性能について調査した結果について述べる。

Performance Evaluation of PC clusters using x86_64 Dual-core CPUs

Tomonori Kouya*

Computer environments for large-scale scientific computation are necessary in many research or educational institutions. In particular, so-called “commodity PC clusters” that are constructed with consumer PCs are widely used in university laboratories. This is because these PC clusters are useful for both research and educational purposes. In 2005, Intel and AMD Corp. released two x86_64 dual-core CPUs; that are compatible with the current x86_32 CPUs and can provide a 64bit environment. Although this will drive an expansion in the capabilities of the large-scale parallel computations in these laboratories, not many publications exist on the performance evaluation of IEEE754 double precision, multiple precision numerical computation and network of these new CPUs. In this paper, we describe the results of these performance evaluations of the Pentium D 820 and the AMD Athlon64 X2 3800+; x86_64 dual-core CPUs with regard to consumer PCs.

1. 初めに

大規模な科学技術計算を実行するための計算機環境は、さまざまな所で求められるようになってきているが、専用ラックに収まった形態のサーバマシンは並列及び分散処理専用であり、購入においても保守においてもコンシューマ向け PC cluster に比べるとコストが高くつく。大学研究室レベルにおいては、多少性能においては劣っていても、教育用にも使用できる汎用的なコンシューマ向け PC を組み合わせたコモディティ PC cluster を使用せざるを得ない場合が多い。これであれば、面倒な作業を厭わなければ高額の保守費用を負担せず、パーツ単位の交換とソフトウェアのアップデートを自分で行うことで維持することが可能となる。

2005 年には現行の x86_32 CPU と互換性を持ち、なおかつ 64bit 環境を提供できる Dual-Core CPU が AMD 及び Intel から登場し、コモディティ PC cluster でもより大規模な並列計算が実行できるようになりつつある。しかし、IEEE754 倍精度・多倍長数値計算及びその並列化に関するこれらの CPU 性能を調査した文献はまだ少ない。今回我々は、コンシューマ向け PC 用の x86_64 Dual-core CPU である Intel Pentium D 820 と AMD Athlon64 X2 3800+ を搭載したマシンを 2 台揃え、これらの数値計算性能及び

ネットワーク性能について調査した。本稿ではその結果について述べる。

2. 実験環境について

使用する 64bit 環境とその略称を表 1 に示す。また比較のために表 2 に示す 32bit 環境も適宜使用する。

表 1: x86_64 Computation Environments

Abbreviation	CPU/OS/Compiler/Mother Board
PentiumD	Intel Pentium D 820 (2.8GHz/Dual-core/L2 2x1MB) Fedora Core 4 x86_64 (Kernel Ver.2.6.14-1) gcc 4.0.2 (w/o any options) ASUS P5WD2 Premium(955X)
Athlon64X2	AMD Athlon64 X2 3800+ (2.0GHz/Dual-core/AMD64/L2 512KB) Fedora Core 4 x86_64 (Kernel Ver.2.6.14-1) gcc 4.0.2 ASUS A8N-SLI Deluxe(nForce4)

*静岡理科大学

*Shizuoka Institute of Science and Technology

多倍長浮動小数点ライブラリは MPFR/GMP を使

表 2: x86_32 Computation Environments

Abbreviation	CPU/OS/Compiler/Mother Board
Pentium4	Intel Pentium 4 2.8cGHz Vine Linux 3.1 (Kernel Ver.2.4.27-0) gcc 3.4.3 (w/o any options)
Xeon	Xeon 3.0GHz (2 SMP) Redhat 8 (Kernel Ver.2.4.20-18) gcc 3.2

用した。GMP は全て Version 4.1.4 を使用し、MPFR は Pentium4 でのみ Version 2.1.1、他は全て Version 2.2.0 を使用した。

3. Serial Computing の性能評価

n 次実正方行列 A, B はどちらも次のようにして生成したものを使用した。

1. rand() 関数を用いて $\hat{A} = [\hat{a}_{ij}]$, $\hat{B} = [\hat{b}_{ij}]$ を生成する。
2. $A = \sqrt{2} \hat{A}$, $B = \sqrt{3} \hat{B}$

こうして生成された行列を用いて、

$$C = AB$$

を計算し、その計算時間と FLOPS 数をカウントする。なお、FLOPS 数は乗算のみ対象とした。通常、行列積によるベンチマークでは加算も対象としているため [4]、ATLAS の Web ページにある timings データと比べ、本ベンチマーク結果の FLOPS 値は 1/2 倍未満に留まっている。

3.1 ATLAS の正方行列積性能

ATLAS[4] は、その名前が示す通り、計算環境に適した行列・ベクトル演算ルーチンを自動的に生成し、高速な基本線型計算を可能にするライブラリである。今回は ATLAS Version 3.7.11 を使用し、全ての環境においてソースからコンパイルしてインストールしたものを使用した。コンパイルに際して使用した arch 名は、Pentium4, Xeon が共に Linux_P4SSE2, PentiumD が Linux_P4E64SSE3, Athlon64X2 が HAMMER64SSE3 である。使用したプログラムでは、Level-3 BLAS[5] の倍精度行列演算 DGEMM サブルーチンを ATLAS が提供する CBLAS インターフェイスから呼び出して行列積 $C = AB$ を計算している。この結果を図 1 に示す。

非常に高速な計算を行うため、低次元 100 次元未満の FLOPS 値が正確に求められておらず (図 1 上)、この辺りのパフォーマンスは比較対象にならない。

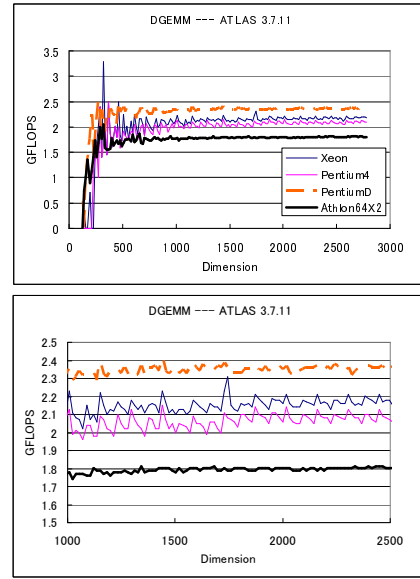


図 1: DGEMM Performance of ATLAS 3.7.11

1000 次元を超えたところを拡大してみると全体としてきれいな水平線が得られており (図 1 下), CPU 性能をよく引き出していることがわかる。この例では PentiumD が最も高速であり、Xeon, Pentium4 がそれに続いて健闘しており、Athlon64X2 が最も性能が悪い。

3.2 BNCpack の正方行列積性能

BNCpack[6] は MPFR と GMP で提供される多倍長浮動小数点演算機能を用いた数値計算ライブラリである。ここでは BNCpack がサポートしている密行列積関数 mul_dmatrix の性能評価を行う。行列は ATLAS のベンチマークで使用した A, B を使用する。

mul_dmatrix 関数は、単純な三重ループを使って行列積を計算する。従って、ATLAS のような高性能は期待できないが、最低のパフォーマンスを調べる目的には都合が良い。その結果を図 2 に示す。

同 clock、同メーカーにも関わらず、低次元の行列積においては、Pentium4 が PentiumD より性能が良い。しかし、高次元になるとその差は殆どなくなり、むしろ L2 cache size に勝る PentiumD の方が平均的には良くなっている。Athlon64X2 は概して低レベルである。

3.3 MPFR/GMP の基本演算・初等関数性能

GMP[7] は、多倍長整数 (mpz_t 型)、有理数 (mpq_t 型)、浮動小数点数 (mpf_t 型) の 3 種類のデータ型とその基本演算を提供しているライブラリである。これらの多倍長演算は全て自然数ライブラリ (mpn) をベースとして組み上げられたもので、計算環境ごとに異なる高速化技法 (SIMD 命令、各種

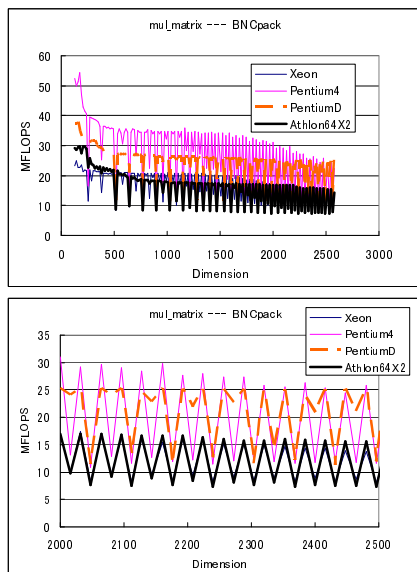


図 2: MUL_MAT Performance of BNCpack

アセンブラルーチンの使用等)は全てこの自然数ライブラリに対して適用されている。自然数ライブラリは整数及び bit 演算によって構築されており、浮動小数点数命令は殆ど使用されていない。

この GMP の自然数ライブラリ的高速性には定評があり、幾つかの多倍長ライブラリにおいても利用されている。MPFR[8] もその一つで、GMP に足りない機能を補った多倍長浮動小数点 (mpfr_t 型) ライブラリである。従って、MPFR のパフォーマンスは GMP の自然数ライブラリのそれに依存して決定されることになる。

MPFR, GMP とも、64bit 環境では仮数部や指数部が 64bit(8 Byte) に拡張される (図 3)。そのため、同じ精度・同じ演算でも、32bit 環境より 64bit 環境の方がループの回数が減る。

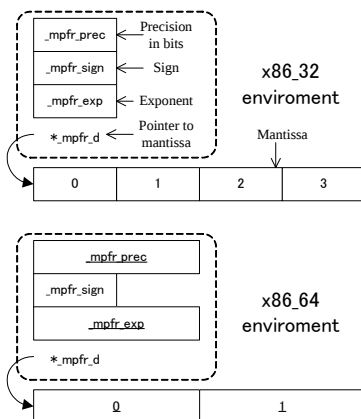


図 3: Changes of mpfr_t data size in 64bit environment

2006 年 1 月上旬現在では、GMP は正式には x86_64 環境向けの高速化技法を行っていないが、64bit 整数演算・bit 演算の速度が 32bit のそれに比べてそれほど落ちていないのであれば、ある程度の性能向上が見込める。

今回は、MPFR Project[8] の timings ページで提供されている timings-mpfr.c を使用して、MPFR の基本演算のうち乗算と除算の性能を測ってみる。実際に行っている計算は

$$x = \sqrt{3} - 1, y = \sqrt{5}$$

とし、乗算は

$$z := xy$$

を、除算は

$$z := x/y$$

である。その結果を図 4 に示す。

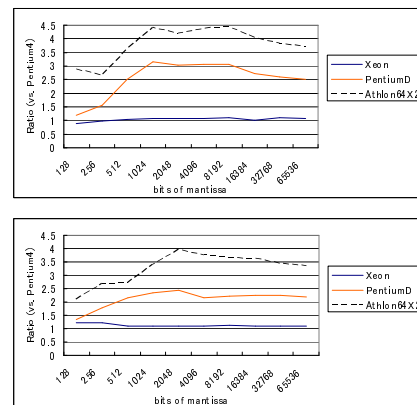


図 4: Performance of MP multiplication and division

32bit 環境では Xeon, Pentium4 と同じパフォーマンスであるが、64bit 環境では PentiumD, Athlon64X2 と性能向上が見られる。対 Pentium4 の比で見ると、Athlon64X2 は、乗算が約 2.5 倍から 4.5 倍、除算が 2 倍から 4 倍程度の性能向上、PentiumD は乗算が 1.2 倍から 3 倍、除算が 1.3 倍から 2.5 倍の性能向上となっている。整数演算とメモリ転送能力 (後述) 能力に秀でた Athlon64X2 の方が、MPFR の多倍長計算では高いパフォーマンスを示している。

4. GbE 性能評価

今回使用した x86_64 環境では、PentiumD, Athlon64X2 どちらも Dual GbE(Gigabit Ethernet) のマザーボードを使用しており、どちらも Intel と Marvell の LAN controller を搭載したものとなっている。Switch も含めた GbE 環境は次の通りである。

PentiumD

eth0 Intel PCI Express Gigabit LAN Controller
eth1 Marvell 88E8001 Gigabit LAN Controller
GbE Switch DELL PowerConnect 2716

Athlon64X2

eth0 NVIDIA nForce4 Gigabit LAN Controller
eth1 Marvell 88E81001 Gigabit LAN Controller
GbE Switch Corega SW05GTV2

今回のベンチマークを行うに当たって Dual GbE の性能比較を行ったところ、Intel Controller の方がほぼ 90% のピーク性能を出しているのに対し、Marvell controller では 55% 程度に留まった。よって、MPI では Intel controller の方を使用している。

ベンチマークには全て NetPIPE 3.6.2[9] をソースからコンパイルしたものを使用している。

4.1 memcpy 性能と node 内の TCP, MPI 性能 まず memcpy 関数の性能を図 5 に示す。

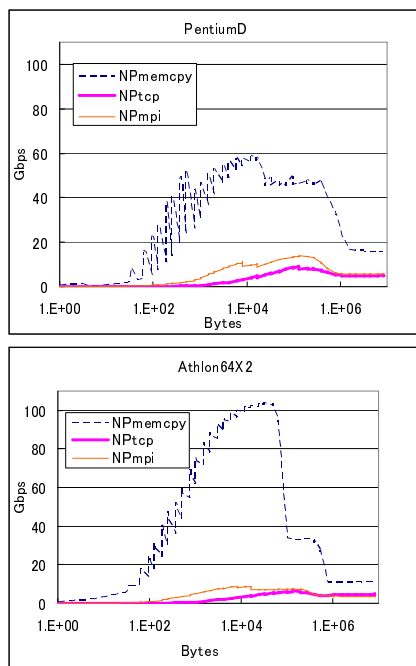


図 5: MEMCPY Performance

最高性能では、Athlon64X2 が 103Gbps、Xeon が 87Gbps、Pentium4 が 78Gbps となっており、Athlon64X2 が最も高性能である。PentiumD はこの中でも最低で 58Gbps にしか達していない。普通に考えれば、この結果は 1 node 内通信性能にもダイレクトに反映される筈であるが、実際にはそうではない。

1 node 内における TCP 及び MPI 通信性能を計測した結果を図 6 に示す。

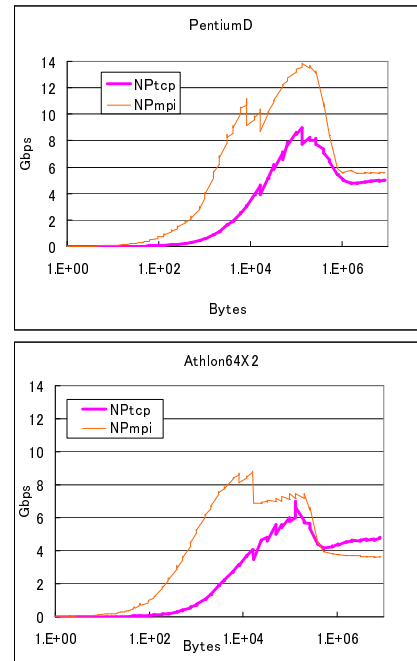


図 6: Performance of TCP and MPI in 1 node

図 5 とは異なり、PentiumD における MPI 通信性能が最も良く、約 14Gbps に達している。それに対し、Athlon64X2 では約 9Gbps、Xeon では 4.3Gbps でしかない。Xeon が低性能なのは、rsh を利用した MPICH Ver.1.2.5 を使用していることも原因と考えられるが、同じ LAM 1.7.7 ベースの RPM パッケージを使っている PentiumD と Athlon64X2 の性能が、memcpy の性能と逆転している理由は現時点では不明である。

4.2 GbE 上の TCP, MPI 性能

我々が以前 Pentium4 で構築した PC cluster では、GbE の性能が十分発揮されていないことが問題点として残っていた [3]。それに対して、SMP マシンである Xeon cluster はほぼ 90% の性能を発揮していたため、この差はソフトウェアと CPU 性能によるものと考えられていたが、それを裏付けるベンチマーク結果を得ることが出来ずにいた。今回 PentiumD と Athlon64X2 cluster を接続する GbE 上において NetPIPE を実行し、ようやくそれを裏付けることが出来た。その結果を図 7 に示す。

Pentium4 では、Windows 上で今回同様、MPICH2 を動作させることにより、TCP 及び MPI 性能を 100Mbps 程度高めることが出来ることは確認できているが、それでも Xeon SMP マシンには遠く及ばなかった。しかし、PentiumD 及び Athlon64X2 では、共に TCP では Xeon SMP 並みの性能を、MPI ではそれを越え、TCP 性能とほぼ同等の性能を得ることが出来ている。

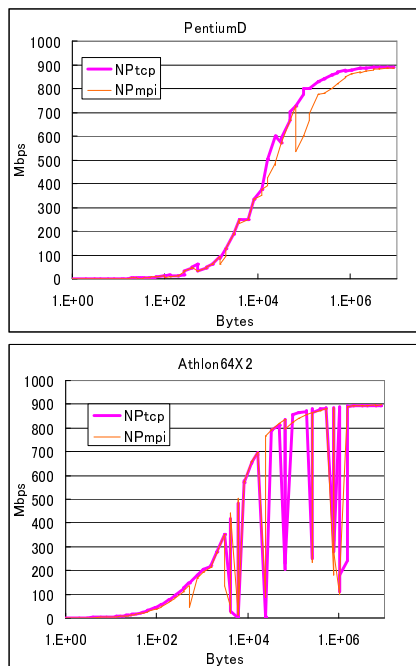


図 7: Performance of TCP and MPI on GbE

しかし、Athlon64X2 マシンではソフトウェアによるものか、nForce4 チップセットによるものかは判然としないが、通信自体が著しく不安定であることが分かる。GbE Switch に問題があるのかと、PentiumD で使用していた DELL の Switch と交換して Athlon64X2 cluster で計測しなおしても、結果に変化はなかった。

4.3 MPIBNCpack の並列正方形行列積性能

以上のベンチマーク結果を踏まえ、目的の並列多倍長数値計算によるベンチマーク結果を見ることにする。図 8 に並列性能を、図 9 に Pentium4 を 1 とした時の性能向上比をグラフにして示す。

並列性能では、短い 128bit 計算では PentiumD, Athlon64X2 のどちらも Pentium4, Xeon に比べて劣る。Node 内通信、GbE 通信とも Pentium4, Xeon より劣っている訳ではないにも関わらず差が出た理由は不明である。Athlon64X2 が PentiumD に劣るのは、恐らく GbE 通信の不安定性によるものと思われる。

全体的なパフォーマンスを比較すると、bit 数の多い計算では PentiumD, Athlon64X2 どちらも安定して Pentium4 の 2 倍以上の性能になることが分かる。逆に bit 数も計算量も小さくなると、Pentium4 との差は縮まってしまう。

PentiumD と Athlon64X2 との比較においては、全体的には乗算性能比と同様、整数演算性能に優る Athlon64X2 が優れている。しかし並列計算では通信の不安定さのためか、その差を縮める結果となった。

5. まとめ

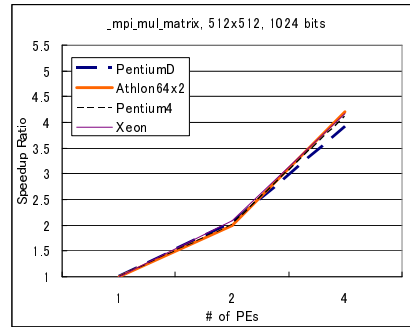
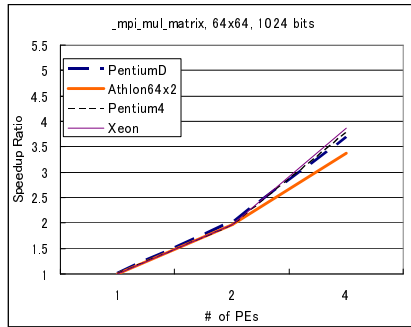
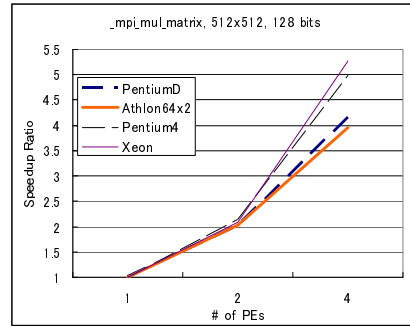
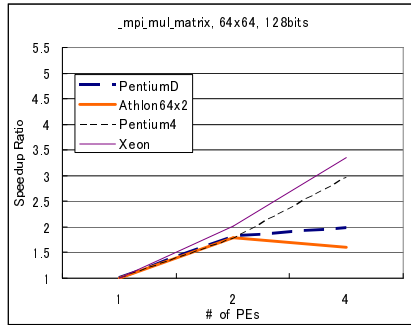
- IEEE754 浮動小数点演算であれば、ATLAS を用いた場合も、単純なループ演算を使用した場合も、PentiumD が Athlon64X2 より優れている。32bit プロセッサとの比較では、ATLAS を用いたときのみ、PentiumD が Pentium4, Xeon を凌駕することが出来た。
- 整数演算主体の多倍長浮動小数点計算なら Athlon64X2 が、PentiumD より優れている。また、bit 数が多くなれば 32bit 環境よりも 64bit 環境の方が 2 倍以上の性能を得られる。
- GbE を用いた場合、ネットワーク性能は CPU 能力に依存して決定されるため、Dual-core CPU は SMP と同等以上の性能が得られ、PentiumD, Athlon64X2 とともに最高性能は 900Mbps に達している。但し、Fedora Core 4 を使用した Athlon64X2 の GbE 通信は著しく不安定である。
- 並列多倍長計算においては、Dual-core CPU は PentiumD も Athlon64X2 も、SMP 並みの並列性能を発揮している。全体的には整数演算主体の多倍長計算に分のある Athlon64X2 が有利となる。

謝辞

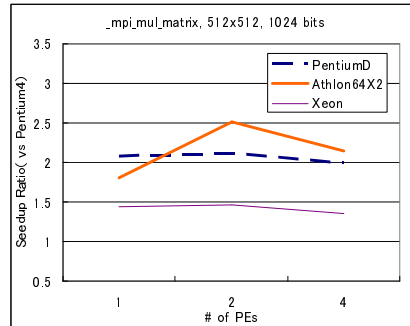
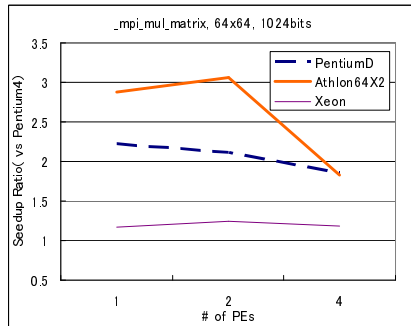
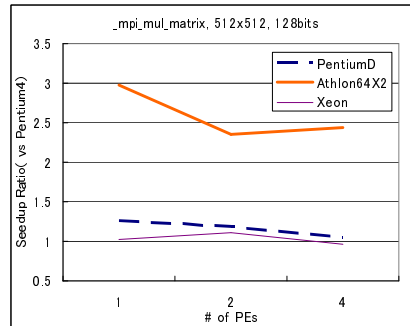
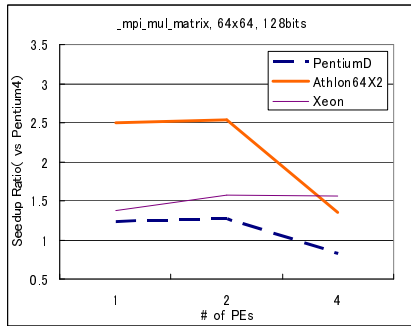
cs-pccluster2 及び Pentium D cluster は静岡理工科大学学内研究費の補助を受けて構築したものである。VTPCC は名古屋大学情報科学研究科三井斌友研究室所有のものを使用させて頂いた。関係者各位に感謝する。

参考文献

- [1] Tomonori Kouya, Performance Evaluation of Multiple Precision Numerical Computation using x86_64 Dual-core CPUs, FCS2005 Poster Session, 2005.
- [2] 幸谷智紀, Windows を用いた PC cluster 上における並列多倍長数値計算ライブラリ MPIBNCpack の性能評価, HPCS2005 ポスターセッション, 2005.
- [3] 幸谷智紀, PC cluster の性能評価 -メモリ及びネットワーク帯域計測編-, <http://na-inet.jp/na/netpipebench.pdf>
- [4] Automatically Tuned Linear Algebra Subroutines, <http://math-atlas.sourceforge.net/>
- [5] Basic Linear Algebra Subprograms, <http://www.netlib.org/blas/>
- [6] BNCpack, <http://na-inet.jp/na/bnc/>
- [7] GMP, <http://swox.com/gmp/>
- [8] MPFR, <http://www.mpfr.org/>
- [9] NetPIPE, <http://www.scl.ameslab.gov/netpipe/>
- [10] LAM/MPI, <http://www.lam-mpi.org/>



⊗ 8: Speedup Ratio of Parallelized MP MatMul



⊗ 9: Speedup Ratio (vs Pentium4)