Vol.2012-HPC-136 No.6
2012/10/3

SR16000/M1とBG/Qにおける性能,精度評価

濱口 信行[1,a)]　石川 正[1,b)]

概要：現在の代表的なスーパーコンピュータ SR16000/M1 と BlueGene/Q において, 所要メモリ量と演算量を変えた演算性能テストプログラム、および幾つかのアプリケーションプログラムを実行した. 両計算機のコアには, SIMD(Single Instruction Multiple Data), SMT(Simultaneous MultiThreading) の機能を有しており、それぞれの特徴を最大限に利用するため、実行性能を評価した。

Evaluation of performance and accuracy on SR16000/M1 and BG/Q

1. はじめに

高エネルギー加速器研究機構のスーパーコンピュータには,SR16000/M1 と BG/Q があり, 性能に関する概要は以下の様になっている.

(1) SR16000/M1(計算機 A) 1 ノード：物理コア 32(論理コア 64). 理論演算性能 980.48GFLOPs. 1 物理コア当たりあたり理論演算性能 30.64GFLOPs.
　L1D キャシュ　32KB コアon chip
　L2 キャシュ　256KB コアon chip
　L3 キャシュ　32MB 8 コアon chip

(2) BG/Q(計算機 B) 1 ノードは物理コア 16(論理コア 64). 理論演算性能 204.8GFLOPs.
　L1D キャシュ　16KB コアon chip
　L2 キャシュ　32MB ノード share

プログラムの演算量は, 使用する計算機のコンパイラの最適化機能,4 倍精度以上の多倍長精度演算, 数学関数計算などに左右されるため, 性能評価での演算量は,SR16000/M1 で-Os オプションで実行した時に採取した性能モニターの値としている. 以後,SR16000/M1 を計算機 A,BG/Q を計算機 B と記述する.

2. SIMD と SMT

両計算機のコアはともに 4 組の乗加算器を有している. この演算器を効率良く使用するために SIMD(Single Instruction Multiple Data),SMT(Simultaneous multithreading) 機能を有している.

SIMD 機能とは拡張した浮動小数点レジスタを使用して, 複数個のデータに対して 1 命令で複数の演算を行うもので, 計算機 A は 2 演算, 計算機 B は 4 演算を行う. 以後 simd と記述する. SMT 機能とは 1 物理コアに対し複数スレッドを生成するもので, 計算機 A は 1,2 スレッド, 計算機 B は 1,2,4 スレッドがある. 以後 smt と記述する.

simd は複数個のデータに対して 1 命令で演算を実行する.

計算機 A では, 複数個のデータに異なる演算をする処理である複素変数乗算,4 倍精度演算には適用できない.

計算機 B では, 複素変数の 2 演算を 1 命令実行できる様に Complex-Simd を実装している. 4 倍精度演算に関しては計算機 A と同様に適用できない.

このため,4 倍精度演算に対しては, 非常に多数のコアに同一の動作をさせて高い性能を達成している専用機がある.[1]

3. 所要メモリ量と演算量

スーパーコンピュータで性能を引き出すには, 並列化, キャシュの有効利用, 通信オーバーヘッド削減などの検討が必要である. その 1 つの指標として, シュミレーションプログラムの所要メモリ量 M と演算量 E の関係が挙げられる. $E = aM^b$ とすると b が大きい程, また b が同じなら a が大きい程性能を引き出す事ができる. 但し a は演算の粒度である. 指数 b の値と代表的な計算は表 1 の様になる.

計算機 A, 計算機 B の特徴で類似点, 相違点をよく示すもの, および精度との関連が深いケースという観点から, 以下の問題を扱った.

1　高エネルギー加速器研究機構
　茨城県つくば市大 1–1
a)　hamagu@post.kek.jp
b)　tadashi.ishikawa@kek.jp

© 2012 Information Processing Society of Japan

表 1 指数 b と代表的計算

b	代表的計算
1	内積計算, 流体計算, 量子色力学計算
1.5	行列積計算, リンパック計算, ルジャンドル陪関数計算 [2]
2 2	重力多体問題計算, 分子動力学計算, 一次元複素フーリエ変換計算 [3]
3 以上	素粒子反応の計算におけるループ積分計算 [4]

- 量子色力学など非常に適用範囲の広い b=1, 内積計算.
- 球座標系での計算に使用され, 精度にも注意が必要な,b=1.5, ルジャンドル陪関数計算と, メモリアクセスパターンが性能に大きな影響を与える b=2, 一次元複素フーリエ変換計算.
- ループ積分計算から精度との関連が深い b=3,infra box 計算.[5] と, 性能が最も良く出ている b=5, ５次元ループ積分計算.
  尚,infra box 計算 以外はすべて倍精度演算で実行している.

4. 内積計算

内積計算では複素変数 ($a = 8$) の場合 (a,b:複素変数）

```
do i=1,n
sum=sum+a(i)*b(i)
end do
```

と実数変数 ($a = 2$) (a,b:実数変数)

```
do i=1,n
sum=sum+a(i)*b(i)
end do
```

の場合を実行した. また複素変数の場合は演算を
実数変数 ($a = 4$) (ar,br,ai,bi:実数変数)

```
do i=1,n
tr=tr+ar(i)*br(i)-ai(i)*bi(i)
ti=ti+ar(i)*bi(i)+ai(i)*br(i)
end do
```

の場合も実行している. ループ長は $2^{20} = 1048576$, 演算量はすべて 8796GFLOP になる様にして測定した結果を表 2 から表 4 に示した.

これらの結果より以下の４つの事が言える.

(１) 性能が最も良いケースでは計算機 A, 計算機 B ともに実行効率は 15%程度.
(２) simd が適用されない場合は smt によるスレッド数を増やすのが効果的.
(３) simd が適用される場合, 計算機 A は smt=1, 計算機 B は smt=2 または smt=4 が効果的という違いがある.
(４) 計算機 B の Complex-Simd の効果は大きい.

表 2 計算機 A 1 ノード実行時間 (秒) および実行効率 (並列化は自動並列)

a	smt 数	実行時間	実行効率 (%)
8	1	58	15.5
8	2	110	8.2
4	1	65	13.8
4	2	112	8.0
2	1	116	7.7
2	2	128	7.0

表 3 計算機 B 32 ノード実行時間 (秒) および実行効率、並列化はノード間は MPI, ノード内はスレッド並列

a	smt 数	実行時間	実行効率 (%)
8	1	12.7	10.1
8	2	8.4	16.4
8	4	8.9	15.1
4	1	32.5	4.13
4	2	33.4	4.02
4	4	39.1	3.43
2	1	41	3.3
2	2	26	5.2
2	4	23	5.8

表 4 計算機 B a=8,32 ノード実行時間 (秒)

complex simd	smt=1	smt=2	smt=4
使用	12.7	8.4	8.9
未使用	53.2	28.0	17.6
比率	4.2	3.3	2.0

5. ルジャンドル陪関数

ルジャンドル陪関数は,

$$P_n^m(x) = \frac{1}{2^n n!}(1-x^2)^{\frac{m}{2}} \frac{d^{n+m}}{dx^{n+m}}(x^2-1)^n \tag{1}$$

で定義される関数である. 但し m：位数,n:次数,$m > 0$. $m > n$ だと $P_n^m(x) = 0$. $n > m$ に関しては

$$P_{m-1}^m(x) = 0, P_m^m(x) = (2m-1)!!(1-x^2)^{\frac{m}{2}} \tag{2}$$

であり, 漸化式,

$$P_n^m(x) = \frac{2n-1}{n-m} x P_{n-1}^m(x) - \frac{n+m-1}{n-m} P_{n-2}^m(x) \tag{3}$$

で計算出来る.

$$\int_{-1}^{1} (P_n^m(x))^2 dx = 2 \tag{4}$$

となる様に正規化した関数を $Q_n^m(x)$ として, 新しく

$$P_{n,m}(x) \equiv Q_{n+m}^m(x) \tag{5}$$

と定めると,

$$P_{0,m}(x) = \sqrt{2m+1}\sqrt{\frac{(2m-1)!!}{(2m)!!}}(1-x^2)^{\frac{m}{2}} \tag{6}$$

$$P_{1,m}(x) = \sqrt{2m+3}\, x P_{0,m}(x), \tag{7}$$

© 2012 Information Processing Society of Japan

$$P_{n+1,m}(x) = c_{m,n}(xP_{n,m}(x) - d_{m,n}P_{n-1,m}(x)) \quad (8)$$

$$c_{m,n} = \sqrt{\frac{(2n+2m+3)(2n+2m+1)}{(n+1)(n+2m+1)}}$$

$$d_{m,n} = \sqrt{\frac{m(n+2m)}{(2n+2m+1)(2n+2m-1)}}$$

という漸化式となる.

精度チェックとして,$\int_{-1}^{1}(P_{n,m}(x))^2dx$ をルジャンドル-ガウス数値積分法 [6] で計算し誤差を調べている.

$n = m$, ルジャンドル-ガウス積分の分点数を $3m$ として, 誤差調査を実施した結果,$m = 960$ では誤差 10^{-12},$m = 1024$ では誤差 5×10^{-2} と急激に悪化する現象が発生した.

詳細に調査すると,$960 \leq m \leq 1024$ では, $-log_{10}$(誤差(m)) $172 - \frac{m}{6}$ となっている事が判明した.

原因は m が大きい場合,$P_{0,m}$ がアンダーフローにより 0 となり, 積分値に寄与する $P_{n,m}$ の値が 0 になる事による. これは,$m = 1024$ で IEEE754-2008 の 4 倍精度で演算を行い, 結果を倍精度で格納した値で積分値を求めると誤差は 10^{-15} になる事で確認した.

計算機 A, 計算機 B の 4 倍精度での表現可能な数値範囲は倍精度と同じため, この問題の解決法は分点数を増やす事である. $n = m = 1024$, 分点数=16384 で誤差 10^{-12} を得たのでこの条件で性能測定を実施した. 演算量は 1149GFLOP, 並列化は分点に関して行い, 漸化式で計算するため simd は適用出来ず, ノード内のスレッド並列実行も出来ない.

測定条件, 計算機 A:1 ノード自動並列 (smp 数 (スレッド数)=32× smt 数) 計算機 B:32 ノード, フラット MPI(mpi 数=32× 16× smt 数) での実測結果を表 5 に示した.

表 5 ルジャンドル陪関数性能測定結果

計算機	smt 数	性能 (GFLOPs)	実行効率 (%)
A	1	77	7.9
A	2	98	10.0
B	1	162	2.5
B	2	304	4.6
B	4	504	7.7

この結果より, 計算機 A, 計算機 B とも,smt によるスレッド数を増やすのが効果的で, 実行効率もそれぞれ約 10%, 約 8% と大きな差はなく, 計算機固有の違いによる差は出ていない.

6. 一次元複素フーリエ変換計算

一次元複素フーリエ変換 c_l <=> y_k $(l, k = 0, 1, ..., L-1)$ (=> 変換,<= 逆変換)

$$e(x) = e^{2\pi ix} = \cos(2\pi x) + i\sin(2\pi x) \quad (9)$$

$$c_l = \frac{1}{L}\sum_{k=0}^{L-1} y_k e(-\frac{lk}{L}) \quad (10)$$

$$y_k = \frac{1}{L}\sum_{l=0}^{L-1} c_l e(\frac{lk}{L}) \quad (11)$$

この式どうり, すべての l,k に対して複素変数指数関数計算を行う性能測定を実施. 演算量は $204L^2$FLOP.

測定条件, 計算機 A: 1 ノード,64smp(smt=2) の自動並列 計算機 B: 32 ノード,2048mpi(smt=4) のフラット MPI での実測結果を表 6 に示した.

表 6 一次元複素フーリエ変換性能測定結果 (GFLOPs)

L	計算機	性能 (GFLOPs)	実行効率 (%)
2^{15}	A	137	14.0
2^{16}	A	136	13.9
2^{17}	A	137	14.0
2^{18}	A	137	14.0
2^{19}	A	137	14.0
2^{20}	A	137	14.0
2^{15}	B	392	6.0
2^{16}	B	415	6.3
2^{17}	B	428	6.5
2^{18}	B	438	6.7
2^{19}	B	437	6.7
2^{20}	B	438	6.7

複素変数指数関数計算の回数を減らすために,$e(-\frac{l}{L})$ $(l = 0, 1, ..., L-1)$ の値をテーブルにする方法がある.$e(\frac{l}{L})$ の値は $e(-\frac{l}{L})$ の共役複素数なのでテーブルは 1 つで良い.

演算量は $20L^2$ と $\frac{1}{10}$ に減少するが, 間接アドレスメモリ参照が発生する. 計算機 A, 計算機 B とも測定条件はテーブルを使用しない場合と同じにして, テーブル方式の性能向上率を求めた結果を表 7 に示した.

表 7 テーブル方式の性能向上率

L	計算機 A	計算機 B
2^{15}	9.6	2.9
2^{16}	9.5	1.6
2^{17}	8.9	1.4
2^{18}	3.1	1.3
2^{19}	0.23	1.2
2^{20}	0.26	1.3

サイズ 2^{20} だと計算機 B は計算機 A のテーブル使用時の約 15 倍の性能, テーブル未使用時の約 4 倍の性能となっている. テーブルの大きさはサイズ 2^{15} で 0.5MB,2^{17} で 2.0MB,2^{18} で 4.0MB.

計算機 A では,L2 キャシュ 256KB コア, 2MB 8 コアで, サイズ 2^{17} までは L2 キャシュアクセス, サイズ 2^{18} 以上では L3 キャシュアクセスとなる.

計算機 B では, フラット MPI で実行しているため, サイ

ズ 2^{15} で 32MB ノードなので L2 キャシュアクセス, サイズ 2^{16} 以上では主記憶メモリアクセスとなる．ただ,$b=2$ の関係より, 主記憶メモリアクセスでも性能向上率は 1 以下に (性能が悪く) なる事はない. 以上から 2 つの計算機には次の差が出ていると言える.

- 計算機 A は, 複素変数指数関数計算の負荷より, キャシュアクセスの負荷が大きい.
- 計算機 B は, 複素変数指数関数計算の負荷が, キャシュ, 主記憶メモリアクセスの負荷より大きい.

7. infra box 計算

infra box 計算とは以下の数式で表される,3 次元数値積分である.

$$I = \int_0^1 dx \int_0^{1-x} dy \int_0^{1-x-y} dz \frac{1}{D^2}, \tag{12}$$
$$D = -xys - tz(1-x-y-z) + (x+y)\lambda^2 + (1-x-y-z)(1-x-y)m_e^2 + z(1-x-y)m_f^2$$

$s<0, t<0, -s, -t >> m_e >> \lambda$ の場合, 解析近似解は

$$I = \frac{1}{-s(-t+m_f^2)} \ln(\frac{-s}{\lambda^2}) \ln \frac{(-t+m_f^2)^2}{m_e^2 m_f^2} \tag{13}$$

結果の検証には,$s=-500^2$, $t=-150^2$, m_e=0.0005, m_f=150 と通常 $\lambda=10^{-30}$ が使用される． λ(仮想光子)=10^{-30} では, 精度 10 進 6 桁の計算結果を得るには,4 倍精度以上の演算精度が必要で, 倍精度演算では $\lambda=10^{-12}$ 以上でなければならない. また $\lambda=10^{-30}$ より小さい場合も考慮し, 6 倍精度,8 倍精度での演算も試みた.

計算機 A, 計算機 B で,$\lambda=10^{-12}$ で倍精度演算,$\lambda=10^{-30}$ で 4 倍精度, 6 倍精度,8 倍精度での性能測定を実施した.

数値積分計算法は二重指数関数型積分法 [7] を使用している.

テーブルサイズは分点と重み係数を格納する領域のおおきさをあらわしている.

7.1 倍精度

サイズ $n=8192$, 演算量 10447GFLOP, テーブルサイズ 128KB

計算機 A:1 ノード　自動並列

計算機 B:32 ノード　ノード間 MPI, ノード内　スレッド並列

性能測定結果を表 8 に示した.

simd が適用されているため, 実行効率は 27% 前後と高くなっている．smt のスレッド数の増加に対する効果は計算機 A は効果なし, 計算機 B は効果ありと違いがでている.

表 8　倍精度性能測定結果 (GFLOPs)

計算機	smt 数	性能	実行効率 (%)
A	1	268	27.3
A	2	263	26.8
B	1	846	12.9
B	2	1493	22.8
B	4	1766	26.9

7.2 4 倍精度

サイズ n=2048, 演算量 618GFLOP, テーブルサイズ 32KB

計算機 A:4,8 ノード　フラット MPI

計算機 B:32 ノード　フラット MPI

性能測定結果を表 9 に示した.

表 9　4 倍精度性能測定結果 (GFLOPs)

計算機	ノード数	smt 数	性能	実行効率 (%)
A	4	1	248	6.3
A	4	2	418	10.7
A	8	1	495	6.3
A	8	2	836	10.7
B	32	1	253	3.9
B	32	2	475	7.2
B	32	4	705	10.8

計算機 A, 計算機 B は simd が適用されないので,smt のスレッド数の増加に対する効果も同様なものとなり, 実行効率も 10% と計算機固有の差はない.

比較のため, 論理演算性能 1088GFLOPs のグラフィックボード (HD5870　2 ボード) では 290GFLOPs と実行効率は 27% となっていて, 多数のコアを使用する方式の優位性が見てとれた.

7.3 6 倍精度,8 倍精度

サイズ n=2048 演算量 6 倍精度　15127GFLOP, 8 倍精度　32831GFLOP

テーブルサイズ 6 倍精度　96KB,8 倍精度　128KB

計算機 A　1 ノード　32MPI(smt=1),64MPI(smt=2)

計算機 B　32 ノード　フラット MPI

性能測定結果を表 10 に示した.

simd が適用されないため smt のスレッド数増加の効果が出ている.

実行効率 10% 以下 (計算機 A, 計算機 B の smt=1) の場合には $E=aM^b$ の a の差 (6 倍精度, 8 倍精度) が性能の差に出ている.

実行効率 15% 前後なると,a の差 (6 倍精度, 8 倍精度) が性能の差をもたらさない結果となっている.

8. 5 次元ループ積分計算

測定した積分の式は下記のもので, 倍精度演算で解析近

表 10　6 倍精度,8 倍精度性能測定結果 (GFLOPs)

計算機	smt 数	精度	性能	実行効率 (%)
A	1	6 倍精度	90	9.1
A	1	8 倍精度	105	10.7
A	2	6 倍精度	135	13.9
A	2	8 倍精度	142	14.5
B	1	6 倍精度	476	7.3
B	1	8 倍精度	609	9.3
B	2	6 倍精度	883	13.5
B	2	8 倍精度	1013	15.5
B	4	6 倍精度	1197	18.3
B	4	8 倍精度	1198	18.3

似解=0.2762092253588 と全桁一致している.[8] 数値積分計算法は二重指数関数型積分法 [7] を使用している．テーブルサイズは分点と重み係数を格納する領域のおおきさをあらわしている.

$$
\begin{aligned}
I &= \int_0^1 \prod_{i=1}^{6} dx_i \frac{\delta(1-\sum_{i=1}^{6} x_i)}{D^2}, \qquad (14)\\
D &= -x_1^2x_2 - x_1^2x_3 - x_1^2x_4 - x_1^2x_6 - x_1x_2^2 - x_1x_2x_3\\
&\quad -2x_1x_2x_4 - x_1x_2x_5 - x_1x_2x_6\\
&\quad -x_1x_3^2 - 2x_1x_3x_4 - x_1x_3x_5 - x_1x_3x_6 - x_1x_4^2\\
&\quad -x_1x_4x_5 - 2x_1x_4x_6 - x_1x_5x_6\\
&\quad -x_1x_6^2 - x_2^2x_4 - x_2^2x_5 - x_2x_3x_4 - x_2x_3x_5\\
&\quad -x_2x_4^2 - 2x_2x_4x_5 - x_2x_4x_6 - x_2x_5^2\\
&\quad -x_2x_5x_6 - x_3^2x_4 - x_3^2x_5 - x_3x_4^2 - 2x_3x_4x_5\\
&\quad -x_3x_4x_6 - x_3x_5^2 - x_3x_5x_6 - x_4^2x_5\\
&\quad -x_4^2x_6 - x_4x_5^2 - 3x_4x_5x_6 - x_4x_6^2 - x_5^2x_6 - x_5x_6^2,\\
x_6 &= 1 - x_1 - x_2 - x_3 - x_4 - x_5.
\end{aligned}
$$

サイズ $n = 128$, 演算量 2668GFLOP, テーブルサイズ 2KB 計算機 A　 1 ノード　 64smp の自動並列で 352GFLOPs, 実行効率 36% 計算機 B　 32 ノード　2048MPI(smt=4) のフラット MPI で 2639GFLOPs, 実行効率 40% と非常に高い性能が得られた.

計算機 A, 計算機 B の値は, 論理演算性能 1088GFLOPs のグラフィックボード (HD5870　 2 ボード) で 456GFLOPs, 実行効率 42% に匹敵している.

これは, 式が長く simd は適用されないがテーブルサイズが 2KB, ノード内 128KB とキャシュを有効に使用している事と smt のスレッド数増加の相乗効果による.

9. まとめ

スーパーコンピュータの性能を, 少し視点を変えて所要メモリ量 (M) と演算量 (E) との関係 $E = aM^b$ から考察して, 以下の知見を得た.

- SR16000/M1,BG/Q ともに性能向上に SIMD 機構の適用が効果的. 特に BG/Q では Complex-Simd 機構の適用の効果が大きい.
- SIMD 機構が適用される場合,SMT 機構の効果は,SR16000/M1 では小さく, 逆効果の場合もあるのに対し,BG/Q では常に大きいという差がある.
- SIMD 機構が適用されない場合,SMT 機構の効果は SR16000/M1 と BG/Q ともに大きい.
- b が 2 以上の場合は,SMT 機構の効果が大きく, キャシュの有効利用が重要になる.

現在,SR16000/M1,BG/Q が高エネルギー加速器研究機構に導入されてまだ約 1 年, 約半年しか経っていないので, 両計算機の特徴の把握が充分になされているとは言えない. 今後は, 更なる性能評価, より一般的な問題への適用, 他の特徴を持つ計算機との比較に取り組み, その成果を, 研究会や, 高性能計算の扉 [9] などで公開していく予定にしている.

謝辞　本研究は　 HPCI 戦略プログラム分野 5「物質と宇宙の起源と構造」, JSPS 科研費 23540328 の助成を受けたものである.

参考文献

[1] 小池寿紀, 近匡, 石川正, 湯浅富久子, 濱口信行: グラフィックボードを用いたファインマン-ループ積分の高速化 (その 2) 日本物理学会第６６回年次大会 (2010).
[2] 森口繁一, 宇田川　久, 一松信: 岩波数学公式集　（新装版）　特殊函数, 岩波書店 (2005).
[3] 一松信:初等関数の数値計算, 教育出版 (1984).
[4] E. de Doncker, Y. Shimizu, J. Fujimoto, F. Yuasa, Computation of loop integralsusing extrapolation, Comput. Phys. Comm. 159, 145 (2004) .
F. Yuasa, E. de Doncker, N. Hamaguchi, T. Ishikawa, K. Kato, Y. Kurihara, J. Fujimoto,Y. Shimizu, Numerical computation of two-loop box diagrams with masses, Comput. Phys. Comm. 183, 2136 (2012) .
[5] 濱口信行:数値解析からファインマンループ積分の特徴と多倍長精度積分の適用, 京都大学数理解析研究所 数理解析研究所講究録 1791 (2012).
[6] 山内次郎 , 宇野利雄, 一松信 ：電子計算機のための数値計算法　， 培風館 (1972).
[7] 杉原正顯, 室田一雄:数値計算法の数理, 岩波書店 (2003).
H. Takahasi, M. Mori, Double exponential formulas for numerical integration, RIMS Kyoto Univ. 9,721(1974).
[8] S. Laporta, High-precison calculation of multi-loop Feynman integrals by difference equations, Int. J. Mod. Phys. A, 15, 5087 (2000).
F. Yuasa, N. Hamaguchi, Y.Shimizu, E. de Doncker, K.Kato, T.Ishikawa, Progress on the Direct Computation Method, http://indico.cern.ch/contributionDisplay.py?contribId=72&sessionId=13&confId=93877
[9] http://www.jicfus.jp/field5/jp/promotion/hpcdoor