

Fortran におけるファイル入出力の高速化

東京大学情報基盤センター

黒田 久泰

配列データの中身をそのままファイルに保存するような場合は、書式なし入出力を使うことでファイル入出力にかかる時間を短縮することができます。また、書式なし入出力ではバッファレス入出力を使うとさらに高速化されます。

1. 書式付き入出力

書式付き入出力では、**FORMAT** 文を用いて書式付きでファイルの保存を行います。例えば、配列 **a** 全体をファイルに保存する場合は次のようにします。

(書式付き入出力)

```
program main
real*8 a(4000000)
open(10, file=' a. dat' )
write(10, *) a
close(10)
end
```

数値 10 の部分は装置番号を指定します。装置番号は 0 から 2, 147, 483, 647 ($=2^{31}-1$) の間から任意に指定できますが、装置番号 0 は標準エラー出力(**stderr**)、装置番号 5 は標準入力(**stdin**)、装置番号 6 は標準出力(**stdout**)に接続されています。なお、実行時オプションとして「-F port(stdunit(0))」を付けた場合は、装置番号 5 は標準入力(**stdin**)、装置番号 6 と 7 はどちらも標準出力(**stdout**)に接続されます。

配列 **a** の一部だけを保存したい場合、例えば **a(101)** から **a(999)** まで 1 つおきに保存したい場合には、「write(10, *) a」の代わりに「write(10, *) (a(i), i=101, 999, 2)」のように記述します。

2. 書式なし入出力

書式なし入出力では、コンピュータの内部形式のまま保存されます。データの変換が行われないことと、書式付き入出力に比べて入出力のサイズが小さくなるためとても効率的です。例えば、配列 **a** 全体をファイルに保存する場合は次のようにします。

(書式なし入出力)

```
program main
real*8 a(4000000)
open(10, file=' a. dat' , form=' unformatted' )
write(10) a
close(10)
end
```

「form=' unformatted'」の部分で「書式なし入出力」であることを指定しています。また、書式なし入出力の場合も「write(10, *) (a(i), i=101, 999, 2)」のように配列の一部だけを保存することができます。

なお、書式なし入出力を指定して出力したファイルを読み込む場合には、同じように書式なし入出力を指定する必要がありますので注意してください。

3. バッファサイズの変更方法

通常、ファイルへの読み書きを行う際には、高速化のため一時的にデータを蓄えておくためのバッファを利用し、そのバッファサイズ単位でファイル入出力を行います。プログラム実行時に Fortran 実行時ライブラリが使用する入出力バッファのサイズを変更することができます。それには、次のように実行プログラム名の後に実行時オプションを付けてバッファサイズを指定します。

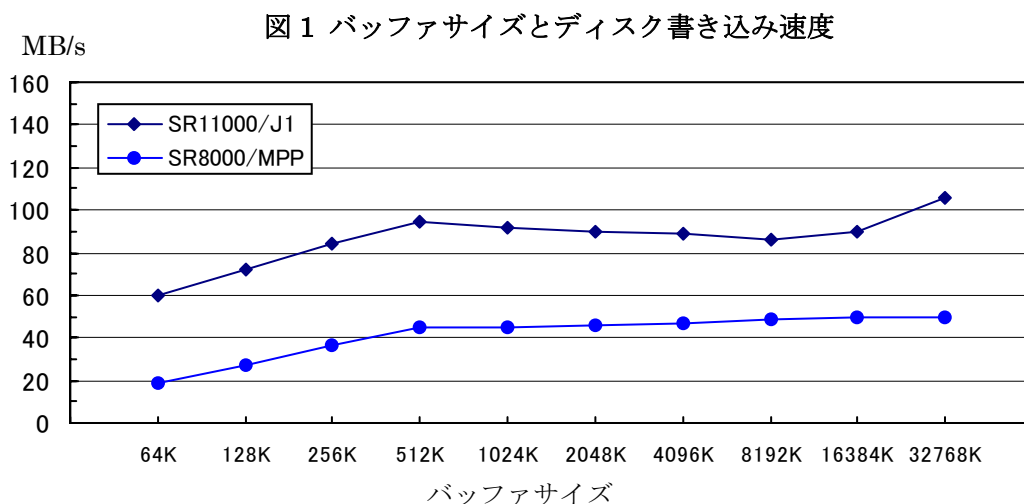
```
% 実行プログラム -F'runst(iobuf(n))'
```

(nに指定した値を8の倍数に切り上げて、キロバイト単位でバッファを確保する)

例：内部バッファサイズを 8192K バイトに設定する場合

```
% ./a.out -F'runst(iobuf(8192))'
```

SR8000/MPP では/home/以下に 1G バイトのデータ、SR11000/J1 では/batch/以下に 2G バイトのデータを「書式なし入出力」で書き込むテストを行った結果を紹介します。バッファサイズを変えたときのディスク書き込み速度は図 1 のようになりました。



Fortran 実行時ライブラリの入出力バッファサイズのデフォルト値は 2006 年 4 月 30 日に従来の 64K バイトから SR8000/MPP では 1024K バイト、SR11000/J1 では 2048K バイトに変更されています。現状ですでに十分なバッファサイズとなっていますので、デフォルト値より大きくしてもあまり性能向上はありません。逆に、バッファサイズを小さくすると性能が低下しますので、ご注意ください。

4. バッファレス入出力

書式なし入出力では、バッファレス入出力を指定することもできます。 バッファレス入出力では、Fortran 実行時ライブラリの入出力バッファを使用しないで直接ファイルとユーザデータ間で入出力を行います。そのため入出力バッファへのデータコピー時間を短縮することができます。

バッファレス入出力に対応させるのはとても簡単です。open 文に「type='bufless'」という TYPE 指定子を付け加えるだけです。また、通常の手式なし入出力では 4GB(=4,294,967,296 バイト)未満のデータしか読み書きできませんが、バッファレス入出力を指定することで 4GB 以上のデータの読み書きもできるようになります。

(バッファレス入出力)

```
program main
real*8 a(4000000)
open(10, file=' a. dat', form=' unformatted', type=' bufless')
write(10) a
close(10)
end
```

なお、バッファレス入出力を指定して出力したファイルを読み込む場合には、同じようにバッファレス入出力を指定する必要がありますので注意してください。

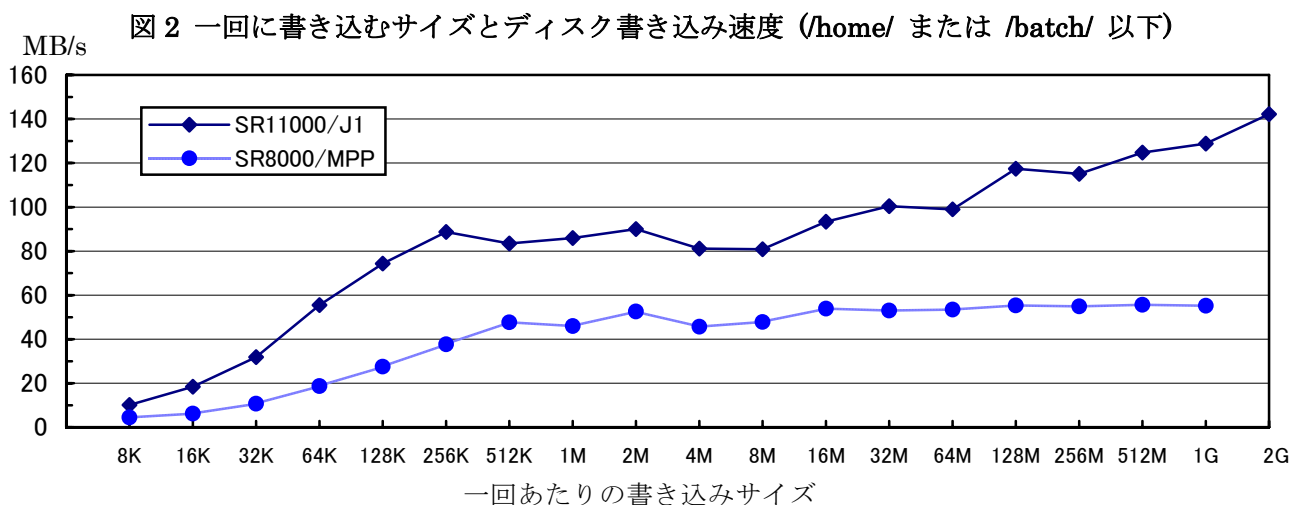
SR8000/MPP では/home/以下に 1G バイトのデータ、SR11000/J1 では/batch/以下に 2G バイトのデータを「書式なし入出力+バッファレス入出力」で書き込むテストを行った結果を紹介します。ここでは、一回に書き込むサイズを変更して測定しています。性能測定プログラムは次のとおりです。

(バッファレス入出力の性能測定プログラム)

```
program main
parameter (MAXSIZE=256*1024*1024)
real*8 a(MAXSIZE), tt
isize=1024
do
  open(10, file=' a. dat', form=' unformatted', type=' bufless')
  call xclock(tt, 7)
  do i=1, MAXSIZE+1-isize, isize
    write(10) (a(j), j=i, i+isize-1)
  end do
  close(10)
  call xclock(tt, 8)
  write(*,*) ' isize=', isize, ' rate=', DBLE (MAXSIZE/131072)/tt
  call unlink(' a. dat')
  if( isize>=MAXSIZE ) exit
  isize=isize*2
end do
end
```

※このプログラムのコンパイルでは、
-64 と -lf90c オプションが必要です

一回に書き込むサイズとディスク書き込み速度の関係は図 2 のようになりました。



SR8000/MPP では 512K バイト、SR11000/J1 では 256K バイトに満たないサイズで何度も書き込むと性能が低下します。また、「write(10) (a(j), j=1, MAXSIZE, 2)」のように飛び飛びに書く場合も著しく性能が低下します。こういった場合は、バッファレス入出力を指定しない方が高速になります。

5. ファイルシステム para-io の利用

単一の大容量ファイルの入出力を行う場合には、ファイルシステム para-io を利用するとさらに高速化されます。「書式なし入出力+バッファレス入出力+ファイルシステム para-io」の組み合わせが最も効果的な利用方法です。ここでは、簡単なプログラム例と NQS スクリプトを示します。

(Fortran プログラム)

```

program main
real*8 a(1024*1024*1024)
open(10, file=' /para-io/*****/a. dat', form=' unformatted', type=' bufless')
write(10) a
close(10)
end

```

※*****/ にはユーザーID を記述
※このプログラムのコンパイルでは-64 が必要です

(NQS スクリプト)

```

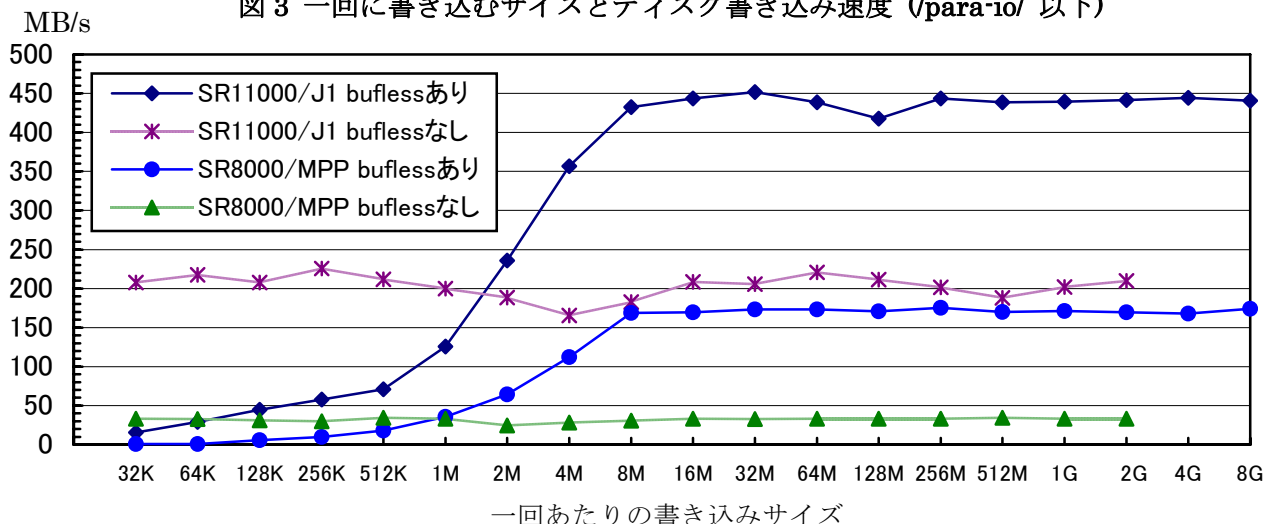
#@$-q personal
#@$-N 1
#@$-lT 00:30:00
./a.out
cp /para-io/*****/a. dat /short/batch/*****/a. dat  ※*****/ にはユーザーID を記述

```

sr11000-s. cc. u-tokyo. ac. jp 上から直接/para-io/以下にアクセスすることはできません。アクセスする場合には、一旦、バッチジョブ上で/short/batch/*****/(*****/はユーザーID)以下にコピーするなどして下さい。/short/batch/*****/ なら sr11000-s. cc. u-tokyo. ac. jp 上からアクセスできます。

前ページの性能測定プログラム中の 'a. dat' を '/para-io/*****/a. dat' に変更すると、ファイルシステム para-io の性能測定を行うことができます。2 行目を parameter(MAXSIZE=1024*1024*1024) にしてファイルサイズを 8G バイトに拡大し、一回に書き込むデータサイズを変更したときの結果を図 3 に示します。また、比較のため「type=' bufless'」を指定しなかった場合も載せてあります。

図 3 一回に書き込むサイズとディスク書き込み速度 (/para-io/ 以下)



バッファレス入出力を指定する場合と指定しない場合でかなり書き込む速度に差が生じることがわかります。バッファレス入出力を指定しない場合には、最高でも SR8000/MPP では約 34MB/s、SR11000/J1 では約 220MB/s までしか書き込み速度が得られません。

一方、バッファレス入出力では一回に書き込むサイズが 8M バイト以上であれば、SR8000/MPP では約 170MB/s、SR11000/J1 では約 440MB/s という十分な書き込み速度が得られます。