

スーパーコンピュータをしよう (2)

永 井 亨

I. はじめに

第1回 [1] では名古屋大学情報連携基盤センターに設置されているスーパーコンピュータ Fujitsu HPC2500 の概要, 利用申請方法, 利用できるソフトウェアについて紹介しました。スーパーコンピュータ (スパコン) は 1664CPU で構成され, 理論最大性能は 13Tflops (1秒間に 13×10^{12} 回の演算ができることとなります), 主記憶 12TB, 磁気ディスク 100TB を搭載しています。大雑把に言って普通の PC (パーソナルコンピュータ) の 500 ~ 1000 倍の規模あるいは処理能力をもつと考えていいでしょう。このシステムは予算上スーパーコンピュータ・システムとアプリケーションサーバ・システムの2つのシステムで構成されていますが, 機種が同一のため一体的に運用されています。このため利用者は実質的に両者を意識する必要はなく, 1つのシステムとして使うことができます。第2回ではスパコンの利用形態とジョブ種別, 利用負担金, 利用例について紹介します。

II. 利用形態とジョブ種別

スパコンの利用形態としては会話型処理 (TSS) とバッチ処理がありますが, いずれの場合も PC 等からネットワークを経由してスパコンに接続 (login) します。ホスト名は hpc.cc.nagoya-u.ac.jp (IP アドレス:133.6.1.153) です。通常, プログラムの編集・翻訳・デバッグ等や対話形式で使用するアプリケーションソフトウェア (SAS, Mathematica, MATLAB など) は TSS で実行します。また, 大規模な計算はジョブを単位としてバッチ処理により実行します。

スパコンで利用可能な CPU 数, CPU 時間や主記憶などの計算機資源の制限値はジョブ種別により異なります。表1に利用できるジョブ種別を示します。必要な資源量に合わせてキューを選択してください。例えば, 4CPU を使い, 各 CPU で CPU 使用時間が5時間かかる並列計算ジョブを実行するには, 4CPU トータルで CPU 使用時間は20時間になるので p8 キューを選ぶのが適当といえるでしょう。主記憶はジョブではなくプロセスごとの制限値になります。したがってこの例では最大 $4 \times 400 = 1600\text{GB}$ が使用できることとなります。もちろん余分に資源を占有することはシステムの効率的運用の妨げになりますので, 必要な資源だけを確保するようご協力をお願いします。

利用者からみて TSS とバッチ処理で同じ点・異なる点を簡単にまとめておきます。

- TSS・バッチ処理ともに同じ実行モジュールを実行しても CPU 使用時間及び経過時間にはばらつきがあります。

- 表1にあるようにバッチ処理の方が計算機資源量が大きいためバッチ処理は大規模計算に向いています。
- TSSでは実行モジュールは起動後、通常ただちに実行が開始されます。バッチ処理ではジョブを投入後直ちに実行が開始されることもあればそうでない場合もあります。他の利用者が計算機資源をすでに使用していて資源の解放を待たなければならないためですが、特にシステムの利用率が高い時期におこります。
- 並列実行ではCPU使用時間から算出される負担経費が異なります。これについてはつぎのセクションで説明します。

表1 ジョブ種別

	キュー名	最大 使用可能 CPU数	CPU使用時間		主記憶		経過時間
			標準値	制限値	標準値	制限値	制限値
バ ッ チ ジ ョ ブ	a8	8	10時間	10時間	2GB	400GB	2時間
	p8	8	10時間	無制限	2GB	400GB	336時間
	p16	16	10時間	無制限	2GB	400GB	336時間
	p64	64	200時間	無制限	2GB	400GB	336時間
	p128	128	200時間	無制限	2GB	400GB	336時間
	p256	256	200時間	無制限	2GB	400GB	336時間
	p1024	1584	200時間	無制限	2GB	400GB	336時間
	TSS	-	64	2時間	無制限	-	128GB

Ⅲ. 利用負担金

第1回でも説明しましたが、スパコンの運営経費の一部は利用負担金として利用者に負担していただいています。まず、利用登録申請1件（アカウント1つ）につき年間10,000円（実際にはこれに5%が上乗せされます）の基本負担金を支払います。これにより年間のCPU使用時間やファイル使用量等に応じた「負担経費」がトータルで20,000円に達するまでシステムを使用できます。要するに1万円前払いすれば2万円分まではスパコンを自由に使えるということなのですが、「負担経費」の算出は少々複雑ですのであとで説明します。

2万円分を使い切るとスパコンの利用はできなくなります。これを回避するため早めに追加負担金の申し込みを行ってください。追加負担金とは不足すると予想される金額を前もって追加するもので、1000円単位で申し込むことができます。追加負担金と利用可能額の関係は表2のようになります。表2は、追加負担金が1万円未満ならばその追加分だけスパコンを利用でき、1万円～10万円未満ならば追加分の2倍まで、...、ということをあらわしていて、追加負担金が大きくなるにつれて利用可能額が優遇されます。なお、支払った追加負担金に対応する利用可

金額に実際に使用した額が達しない場合、その差額は返還されないので注意してください。

表2 追加負担金と利用可能額

追加負担金	追加利用可能額
1万円未満	追加負担金額
1万円～ 10万円未満	追加負担金額の2倍
10万円～ 50万円未満	追加負担金額の3倍
50万円～ 100万円未満	追加負担金額の4倍
100万円以上	追加負担金額の5倍

上で見たように基本負担金・追加負担金ともに前払いとなっていますが、実際にどれだけの計算機資源を使用できるか、あるいは、使用したかは負担経費により算出されます。負担経費には演算負担経費、ファイル使用負担経費、出力負担経費の3つがあります。

●演算負担経費

➤ TSS の場合

CPU 使用時間 1 秒につき 0.1 円です。例えば 20CPU を使用して 1 時間の計算をしたとすると演算負担経費は

$$20 (\text{CPU}) \times 3600 (\text{秒}) \times 0.1 (\text{円 / 秒}) = 7200 (\text{円})$$

となります。実際には各 CPU で CPU 使用時間にはばらつきがありますのでそれぞれの CPU 使用時間をたしあわせたものに 0.1 をかけたものが演算負担経費になります。

➤ バッチ処理の場合

CPU 使用時間 1 秒につき 0.1 円×並列係数で算出されます。並列係数とは使用した CPU 数に応じて決められた係数で表 3 のようになっています。並列効率がよいプログラムが優遇されるようになっていて、例えば 1CPU で計算しても 15CPU で計算しても並列係数は 1 ですから、いずれの場合も CPU 使用時間 1 秒につき 0.1 円ということになります。また、20CPU を使用して 1 時間の計算をしたとすると演算負担経費は

$$3600 (\text{秒}) \times 0.1 (\text{円 / 秒}) \times 2 = 720 (\text{円})$$

となります。これは TSS の場合の 10 分の 1 です！正確にいうとバッチ処理では各 CPU の使用時間の中で最大のもの（各 CPU の使用時間にはばらつきがあるためです）に並列係数をかけて算出します。

したがって、プログラムの編集・翻訳・デバッグ等、CPU 使用時間が比較的小さい作業は TSS で行ない、並列計算はバッチ処理で行うのがスパコンを有効利用するための基本になります。10 万円の追加負担金を支払って 8CPU の並列計算を実行すると、使用可能な時間は最大で

$$220,000 (\text{円}) \div 0.1 (\text{円 / 秒}) = 2,200,000 (\text{秒}) = \text{約 } 25 \text{ 日}$$

になります。

表3 バッチ処理での並列係数

使用 CPU 数	並列係数
1 ~ 15	1
16 ~ 32	2
33 ~ 64	3
65 ~ 128	5
128 ~ 256	10
256 ~ 512	20
513 以上	50

●ファイル使用負担経費

磁気ディスク使用量に応じて発生する負担経費です。この算出方法も少々わかりにくいのですが、表4のようになっています。100MB 以下の使用量ならば負担経費は発生しません。1GB 以下ならば1日あたり20円、1ヶ月あたりに換算すると600円です。500GBでは1日あたり20円 + (500 - 1) GB × 2円 / GB = 1,018円、1ヶ月あたり30,540円です。

表4 ファイル使用負担経費

磁気ディスク使用量	原単位
100MB 以下	0円
100MB を超え 5TB 以下 1GB を超えた部分について	20円 / 日 2円 / GB / 日
5GB を超える場合	15,000円 / 日

●出力負担経費

基盤センター新館1階の利用者端末室に設置されたプリンタに出力する場合に適用される負担経費です。白黒印刷では1ページあたり3円、カラー印刷では1ページあたり8円です。カラープリンタはA0まで印刷できます。

IV. 利用例

1. TSS

スパコンに接続 (login) すると TSS の環境になります。OS は solaris 9 で、emacs と vi が標準のエディタとして用意されています。login 直後の TSS の CPU 使用時間は 120 分に設定されています。これを変更するに limit コマンドを使用します。CPU 使用時間を 5 時間に変更する場合には

```
hpc% limit cputime 5h
```

と入力します。ただし、最初の「hpc%」はシステムから出力されるコマンドプロンプトです。
Cプログラム prog.c を翻訳・実行するには

```
hpc% fcc -o prog prog.c
hpc% ./prog
```

とします。Fortran プログラム・C++ プログラムを翻訳するには、それぞれ firt コマンド・FCC コマンドを使います。詳細は省略しますが、自動並列化機能を使って並列実行するにはオプション -Kparallel を指定して

```
hpc% fcc -Kparallel -o prog_p prog.c
hpc% setenv PARALLEL 8
hpc% ./prog_p
```

とします。PARALLEL は環境変数で 8CPU 使用することを指定しています。

実際の CPU 使用時間は timex コマンドにより知ることができます。

```
hpc% timex ./prog
```

これを実行すると計算結果に続いて以下のようなメッセージが出力されます。

```
real          6:31.22
user          6:23.52
sys           2.11
```

real は経過時間をあらわし、6分 31.22 秒であることを示しています。CPU 使用時間は user と sys の項目を合わせたものです。並列実行の場合には、user と sys の項目は各 CPU の合計になります。timex の情報は標準エラー出力に出力されます。

2. バッチ処理

バッチ処理の手順は以下のようになります。

1. 実行するコマンド（あるいは実行モジュール名など）を記述したスクリプトファイル（バッチリクエスト）を作成する。
2. qsub コマンドでバッチリクエスト（ジョブ）の実行依頼をする。
3. qstat コマンドで依頼したジョブの実行状況の確認をする。
4. ジョブの実行が終了したら出力結果をチェックする。

スクリプトファイル exec.sh の内容は以下のようになります。

```
# @$-q p8 -eo -o prog.out
./prog
```

-q は qsub コマンドのオプションで投入するキュー名を指定します。オプション -eo は標準

出力と標準エラー出力をファイル prog.out に書き出すことを指定しています。つぎに qsub コマンドでバッチ処理を依頼します。

```
hpc% qsub exec.sh
```

計算が実行中かどうかは qstat コマンドで確認してください。

```
hpc% qstat p8
p8@sp0; type=BATCH; [ENABLED, RUNNING]; pri=31
 0 exit;  1 run;  0 queued;  0 wait;  0 hold;  0 arrive;
REQUEST NAME   REQUEST ID   USER        PRI    STATE    PGRP
1: exec.sh     807.sp0     a49999a     31    RUNNING  27001
```

この例ではアカウント a49999a の exec.sh は実行中 (RUNNING) であることが示されています。項目 STATE が QUEUED の場合には実行待ちであることをあらわします。自動並列機能を使って翻訳した実行モジュールをバッチ処理するには以下のようなバッチリクエストを作成します。

```
# @$-q p8 -lp 8 -eo -o exec_p.out
./prog_p
```

オプション -lp は使用 CPU 数を指定します。実際にどれだけの主記憶を使用しているかを知るには qsub コマンドにオプション -oi を指定します。これにより標準出力に以下のような情報が出力されます。Used Resource 欄の Total Max Large Page に示されている値が実際の主記憶使用量です。この値をつぎに実行するときの qsub コマンドのオプション -lM で指定 (5GB 必要ならば 5gb とします) してください。

```
Allocated Resource
Process Resource Set:          1 Set
Total Large Page      :          4098 Mbyte
CPUs                  :           4 CPU
:
Used Resource
Total System CPU Time:         1290 msec
Total User CPU Time   :    32116540 msec
Total Max Large Page  :         1024 Mbyte
```

詳細な利用法は「スーパーコンピュータ及びアプリケーションサーバ利用の手引き」[2]に記載されていますので、実際に利用する場合には是非ご覧になることをおすすめします。

参考文献

- [1] 石井克哉: “スーパーコンピュータをしよう (1)” 名古屋大学情報連携基盤センターニュース, Vol.6, No.3, 263-267 (2007)
- [2] URL: <http://www2.itc.nagoya-u.ac.jp/riyou/tebiki.pdf>

(ながい とおる : 名古屋大学情報連携基盤センター)