

# ベクトル並列型スーパーコンピュータ SR11000/J1 運用開始のお知らせ

システム運用係

スーパーコンピューティングニュース Vol.7 No.2(2005.3)「SR11000/J1 試験運用及びサービス内容の変更について」等にてお知らせしてまいりましたとおり、本センターでは2005年3月にベクトル並列型スーパーコンピュータSR11000/J1を導入し、2005年6月末日まで試験運用、7月から正式運用を開始しました。ここでは本システム運用における、SR8000/MPPを含めたシステム構成、資源設定、利用方法についてお知らせします。

## 1. SR11000/J1 正式運用について

ベクトル並列型スーパーコンピュータSR11000/J1は7月1日より正式運用を開始しました。ノードについては試験運用中と同じく、並列キュー(P001~P004)実行用のノードは物理的な1ノードを2つのノードに論理分割し、8プロセッサで構成するSMPとして運用します。したがって、1ノード当たりの理論演算性能は60.8GFLOPS(SR8000/MPPの約4.2倍)、主記憶容量は64GB(SR8000/MPPの4倍)となります。なお、並列実行ノード数は4ノードまでです。パーソナルコースの方で利用を希望する場合はSR11000オプション(1,2,又は4ノード)をお申込み下さい。

システムは、正式運用当初はバッチシステム環境のみを提供していましたが、8月1日よりインタラクティブ(試用)サービスを開始します。バッチジョブの投入はSR8000/MPP及びSR11000/J1より可能です。ジョブスクリプトのキューの指定にはパイプキュー **single** または **parallel** を使用して下さい。なお、SR8000/MPPとはバイナリ互換性が無いため、SR8000/MPPで使用していたプログラムをSR11000/J1で実行するためには、ソースプログラムの再コンパイルが必要です。

本誌別記事「SR11000/J1におけるインタラクティブ(試用)サービス開始のお知らせ」をご覧ください。

SR11000/J1のソフトウェアは以下のとおりです。

- ・言語プロセッサ(コンパイラ)
  - 最適化 FORTRAN77                      最適化 C
  - 最適化 FORTRAN90                    最適化標準 C++                      国際標準規格 ISO/IEC 14882:1998 準拠 及び  
旧言語仕様 (ARM仕様) 対応
- ・並列化支援アプリケーション
  - MPI
- ・数値計算ライブラリー
  - MSL2                                      MATRIX/MPP/SSS
  - MATRIX/MPP                            BLAS, LAPACK, ScaLAPACK
- ・分子軌道計算プログラム
  - Gaussian03 (Rev C.02)
- 並列実行ランタイム環境 Remote DMA, PVM, 及び並列言語コンパイラ-PARALLELWARE, Parallel FORTRAN は使用できません。
- SR11000/J1システムのハードウェア及びソフトウェアの機能、性能について詳細はスーパーコンピューティングニュース Vol.6 No.6(2004.11)「次期ベクトル並列型スーパーコンピュータシステム SR11000のご紹介」を御覧下さい。

## 2. インタラクティブについて

SR8000/MPP 及び SR11000/J1 を利用するには、以下のホストに SSH (Ver.1 対応) にて接続し、ログインして下さい。

mpp-s.cc.u-tokyo.ac.jp ... SR8000/MPP  
 sr11000-s.cc.u-tokyo.ac.jp ... SR11000/J1 (試用サービス)

sr11000-s.cc.u-tokyo.ac.jp (以下 sr11000-s) は 8 月 1 日より試用サービスを開始します。

- ・ sr11000-s のホームディレクトリーは /batch となります。
- ・ sr11000-s では SR11000/J1 にのみジョブの投入が可能です。SR8000/MPP には投入できません。
- ・ バッチジョブのイニシエーターメールやターミネーターメールの送り先については、ジョブスクリプトに指定のない場合、投入したシステムのアカウントとなります。

詳細は本誌別記事「SR11000/J1 におけるインタラクティブ (試用) サービス開始のお知らせ」を御覧下さい。

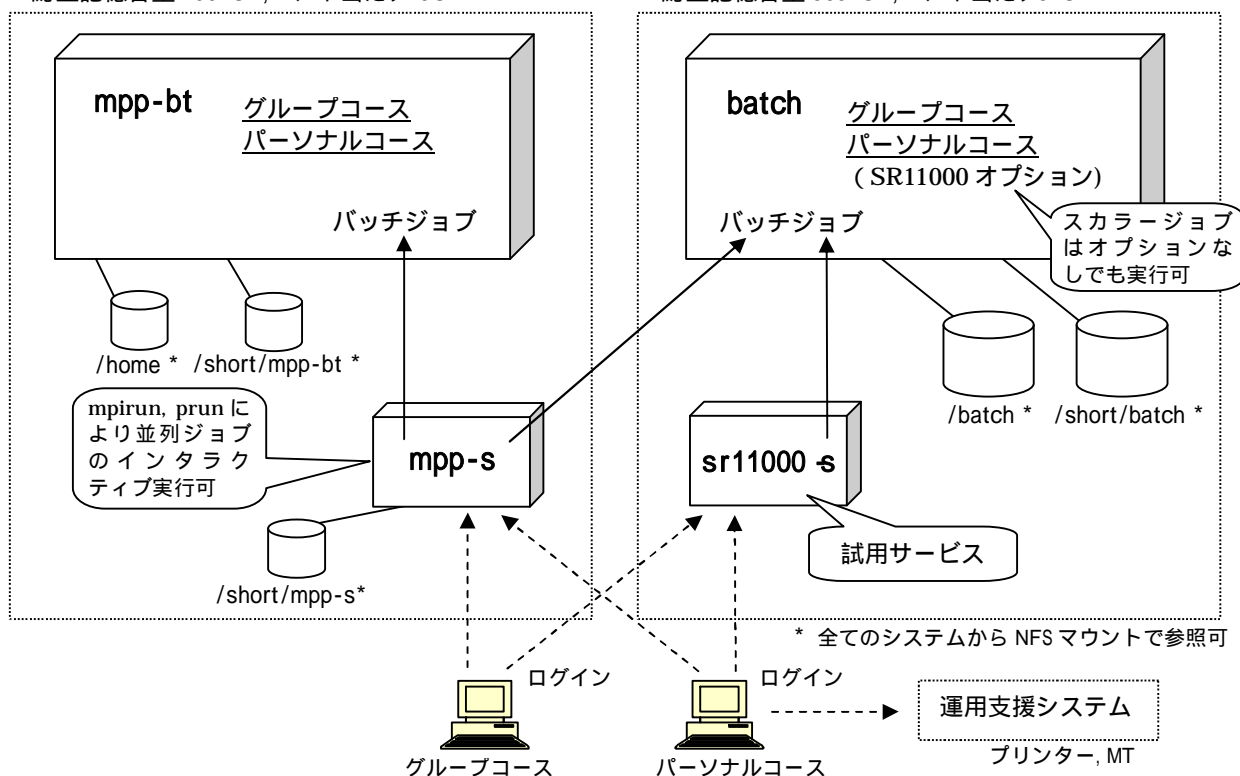
なお、これらのシステムではプログラムのコンパイルが可能です。SR8000/MPP と SR11000/J1 はバイナリー互換ではないため注意が必要です。

ファイル転送については、scp 又は ssh のポート転送機能を利用した ftp を御使用下さい。

### ● システム構成 (2005 年 8 月 ~)

超並列型スーパーコンピューターSR8000/MPP  
 ノード総数 144 ノード  
 総演算性能 2073.6GFLOPS, 1 ノード当たり 14.4GFLOPS  
 総主記憶容量 2304GB, 1 ノード当たり 16GB

ベクトル並列型スーパーコンピューターSR11000/J1  
 ノード総数 88 ノード(1 ノード 8CPU 構成)  
 総演算性能 5350.4GFLOPS, 1 ノード当たり 60.8GFLOPS  
 総主記憶容量 5632GB, 1 ノード当たり 64GB



### 3. バッチジョブについて

バッチジョブは全ての利用者が SR8000/MPP 及び SR11000/J1 のスカラーキュー（A～F）を利用できます。その他のバッチキュー（以下キュー）については、申込のコースにより制限がありますので、次の表で確認して下さい。

コース別バッチキュー利用表 (2005年4月1日変更)

システム	バッチキュー	パーソナルコース			グループコース
		1	2	3	
SR8000/MPP	A～F				
	A-ES～F-ES	×			
	P001				
	P002	×			
	P004	×	4		
	P008	×	8		
	P016	×	×		
	P032	×	×		
	P128	×	×		
SR11000/J1	A～F				
	A-ES～F-ES	×	1	1	
	P001	×	1	1	
	P002	×	×	2	
	P004	×	×	4	

- 注)・ 4, 8 はパーソナルコース 2 において、それぞれ 4 ノード又は 8 ノードを選択した場合に実行可。  
 ・ 1, 2, 4 は SR11000 オプションにおいて、それぞれ 1 ノード, 2 ノード又は 4 ノードを選択した場合に実行可。  
 ・ において、上位のノード数は下位の内容を含む。  
 ・ グループコースは SR8000/MPP においてグループ専用キューを利用可能。

SR8000/MPP と SR11000/J1 とでは、1 ノード当たりの理論演算性能（それぞれ 14.4GFLOPS, 60.8GFLOPS）及び主記憶容量（それぞれ 16GB, 64GB）が異なります。このため、キューの制限時間、使用できるメモリーの上限值にも違いがありますのでジョブクラス制限値（キューの制限値）に御注意下さい。SR11000/J1 のジョブクラス制限値は以下のとおりです。SR8000/MPP のジョブクラス制限値は本誌 2 ページを御覧下さい。

SR11000/J1 ジョブクラス制限値 (2005年3月23日)

キュー名	制限時間 E-TIME	メモリー (GB)		ノード数
		仮想メモリー	区分 ES	
(single)				
A	10 分	2(56)	-	-
B	1 時間	"	-	-
C	5 "	"	-	-
D	15 "	"	-	-
E	50 "	"	-	-
F	225 "	"	-	-
A-ES	10 分	2(56)	0(112)	-
B-ES	1 時間	"	"	-
C-ES	5 "	"	"	-
D-ES	15 "	"	"	-
E-ES	50 "	"	"	-
F-ES	225 "	"	"	-
(parallel)				
P001	64 時間	56(56)	-	1
P002	32 "	"	-	2
P004	16 "	"	-	4

- 注)・ キュー名の single, parallel はパイプキューの名称。  
 ・ CPU 時間 (C-TIME) は制限しない。  
 ・ 仮想メモリーの大きさは 1 ノード当たりの標準値（括弧内は最大値）。  
 ・ 1 ノードは 8CPU 構成。

## ジョブクラスについて

- ・スカラージョブクラス (A, B, C, D, E, F)  
スカラープログラムを実行するジョブクラスです。ジョブの実行時間 (経過時間) でクラス分けしています。
- ・拡張記憶 (ES) 使用ジョブクラス (A-ES, B-ES, C-ES, D-ES, E-ES, F-ES)  
拡張記憶 (ES) を使用するジョブ実行のためのジョブクラスです。ジョブの実行時間 (経過時間) でクラス分けしています。

拡張記憶 (ES) とはメモリーを外部記憶装置とみなして入出力を行う大容量擬似ファイルシステムです。例えば次のようにファイルとして宣言して利用します。

```
open(60,file='esdata',type='es',space='10GB',form='unformatted')
```

- ・並列ジョブクラス (P001, P002, P004)  
要素並列プログラムの実行及び複数ノードを使用したジョブの実行を行うジョブクラスです。使用ノード数と実行時間 (経過時間) でクラス分けしています。

各キューへのジョブの投入にはパイプキューを指定して下さい。ジョブが投入されるキューは制限時間やノード数の制限値の指定により自動的に選択されます。パイプキューは SR11000/J1 では **single** 又は **parallel**, SR8000/MPP では **mpp-single** 又は **mpp-parallel** を指定します。バッチジョブのスキプットの記述は以下の例を参考にして下さい。なお、例ではキューの選択に必要なオプションのみ記載しており、その他の記述は省略していますので実際にスクリプトファイルを記述する場合には適宜補って下さい。

### SR8000/MPP のキューA~F

```
#@$ -q mpp single
#@$ -IT 10:00:00 キューD
#@$ -IM 2GB
```

### SR11000/J1 のキューA~F

```
#@$ -q single
#@$ -IT 5:00:00 キューC
#@$ -IM 2GB
```

### SR8000/MPP のキューA-ES~F-ES

```
#@$ -q mpp single
#@$ -IT 5:00:00 キューC-ES
#@$ -IV 1GB
```

### SR11000/J1 のキューA-ES~F-ES

```
#@$ -q single
#@$ -IT 10:00:00 キューD-ES
#@$ -IV 50GB
```

### SR8000/MPP のキューP001~P032

```
#@$ -q mpp parallel
#@$ -N 1 キューP001
#@$ -IT 20:00:00
```

### SR11000/J1 のキューP001~P004

```
#@$ -q parallel
#@$ -N 4 キューP004
#@$ -IT 10:00:00
```

### SR8000/MPP のキューP128

```
#@$ -q mpp parallel
#@$ -N 128 キューP128
#@$ -IT 1:00:00
```

月 1 回の 128 ノードジョブ実行サービス時のみ。

グループコースの利用者は上記のキューに加え SR8000/MPP の専用キューを使用することができます。

### SR8000/MPP のグループ専用キュー

```
#@$ -q a00
#@$ -N 1 キュー-a00
#@$ -IT 24:00:00
```

キューの制限時間は最大 24 時間、メモリーは 1 ノード当たり最大 13.5 GB。登録ノード数まで利用可。  
(グループコードが a00 の場合の例)

なお、SR8000/MPP と SR11000/J1 はバイナリー互換では無いため、一方でコンパイルしたプログラムを他方で実行することはできません。それぞれのシステムにおいてコンパイルを行う必要があります。

#### 4. ディスクについて

利用者ディスクとして長期保存ファイル用に /home, /batch, 短期保存ファイル用に /short/mpp-bt, /short/batch, 並列入出力用に /para-io の各ディレクトリーを用意しています。各コースで使用できるディレクトリーは以下のとおりです。

(2005年4月1日変更)

システム	ディレクトリー	保存日数	パーソナルコース			グループコース
			1	2	3	
SR8000/MPP	/home	-				
	/short/mpp-s	5日間				
	/short/mpp-bt	15日間				
	/para-io	5日間				
SR11000/J1	/batch	-				
	/short/batch	15日間				
	/para-io	5日間				

注)・ はファイル上限値の設定を行わない。

・実際に利用できるディレクトリーは、上記の各ディレクトリー下の利用者番号名ディレクトリーとなります。  
例: /batch/p05000/

各ディレクトリーは /para-io を除きシステム間で NFS マウントしますので、SR8000/MPP, SR11000/J1 のどちらのシステムからも参照できます。ただし、一方のシステムが停止した場合には他方から参照できなくなりますので、できるだけローカル接続（NFSではなく、システムに直結）されたディスクを使用することをお勧めします。

また、/home, /short/mpp-bt, /batch, /short/batch の各ディスク容量上限値はパーソナルコースは利用者ごと、グループコースはグループ全体に設定され、設定値は申込み時の「ディスク量」（la-d コマンドで確認可能）です。

/short/mpp-s 及び /para-io は申込みの必要はありませんが、利用者ディレクトリーとして利用することができます。これらのディレクトリーは保存期間を5日間としますが、ファイル上限値を設定しませんので、利用状況によっては容量不足によりファイルが作成できないことも起こり得ます。むやみに長期間使用しない、使用後はすぐに削除する等、節度ある利用を心掛けて下さい。なお、慢性的に容量不足となる場合は利用を制限することがありますので御了承下さい。

#### 5. SR11000/J1 正式運用における注意事項について

SR11000/J1 をご利用の際は、次の点にご注意下さい。

- ・バッチジョブを投入する際は qsub コマンド (/usr/ucb/qsub) を使用して下さい。
- ・ユーザーの投入可能ジョブ数は8、実行可能ジョブ数はキュー毎には1、キューが異なれば2となります。
- ・バッチジョブの削除は、ホスト名まで指定して下さい。

例)

qdel リクエスト番号.ホスト名

SR8000/MPP のバッチジョブで、ジョブを投入したホストから削除を行う場合は ".ホスト名" を省略できます。

- ・ジョブ終了時の標準エラー出力（例: job.csh.e9999）が空の場合、ファイルが作成されません（SR8000/MPP では空の場合でも作成されます）。なお、標準出力（例: job.csh.o9999）は空の場合でも作成されます。

- mpirun 及び prun について  
 mpirun 及び prun のノード数は、ジョブスクリプトで指定したノード数を使用します。-n オプションは無効となります。  
 mpirun のプロセス数はジョブスクリプトのノード数とジョブタイプで決定されます。ジョブタイプに SS を指定すると、プロセス数はノード数の 8 倍となります。ジョブタイプを指定しない場合はノード数と同じ値となります。-np オプションは無効となります。

例 1)

```
#@$-N 4
mpirun ./a.out      4 ノード 4 プロセスで実行します。
```

例 2)

```
#@$-N 4
#@$-J SS
mpirun ./a.out      4 ノード 32 プロセスで実行します。
```

- MPI を使用したプログラムのコンパイルについて  
 MPI を使用したプログラムをコンパイルする場合は、以下のコマンドを使用して下さい。

```
mpif77    (最適化 FORTRAN77)
mpif90    (最適化 FORTRAN90)
mpicc     (最適化 C)
mpiCC     (最適化標準 C++)
```

- C++ における STL (標準テンプレートライブラリー) の使用について  
 次のように指定して下さい。

```
sCC sample.cpp -I/opt/STLport-4.5/stlport -L/opt/STLport-4.5/lib -lstdport_sCC -lm
```

- 擬似ベクトル化について

SR8000/MPP ではソフトウェア (コンパイラーが生成したオブジェクトプログラム) が主体となり擬似ベクトル化を行います。SR11000/J1 には擬似ベクトル化という概念はありません (SR11000/J1 では、プリロード機能が廃止され、新たにハードウェアプリフェッチ機能が追加されるなど動作が異なります)。したがって SR11000/J1 では擬似ベクトル化に関するコンパイラーのメッセージは出力されず、擬似ベクトル化オプション (-pvec) は無効となります。

- pmpr コマンド (性能モニタの結果ファイルの表示) について

pmpr コマンドを実行するシステムと、結果ファイルを出力したシステムが異なる場合においても表示が可能となりました。

- コンパイラーの -model オプションについて

マニュアルでは -model = H1 がデフォルト値と記述されていますが、本センターでは -model=J1 をデフォルト値としています。コンパイル時には特に明示する必要はありません。

- Gaussian03 について

7 月からパーソナルコースユーザーに加え、グループコースユーザーの利用も可能になりました。

- SR11000/J1 の言語設定について

言語設定の標準値は英語です。バッチジョブからの出力を変更する場合は以下を指定して下さい。

csh の場合

```
setenv LANG Ja_JP    (Shift-JIS で出力)
setenv LANG C        (英語で出力)
```

sh の場合

```
LANG=Ja_JP ; export LANG (Shift-JIS で出力)
LANG=C ; export LANG     (英語で出力)
```

コンパイラーのメッセージのみ変更する場合は、以下のコンパイルオプションを指定して下さい。

-listlang = sjis                   (Shift-JIS で出力)  
-listlang = c                   (英語で出力)

- BLAS, LAPACK の利用について  
8 月 1 日より利用が可能となります。詳細は本誌別記事「SR11000/J1 BLAS・LAPACK の利用について」を御覧下さい。

次の点については調整中ですので、今しばらくお待ち下さい。

- ScaLAPACK の利用
- prun コマンドの -f オプションの使用