

# ジョブスケジューリングシステムについて

システム運用係

本センターでは利用者の皆様に有効かつ公平な計算機資源の提供を行うため、超並列型スーパーコンピュータSR8000/MPP 及びベクトル並列型スーパーコンピュータSR11000/J1 のバッチジョブの実行においてジョブスケジューリングシステムを導入しています。詳細は以下のとおりです。

本ジョブスケジューリングシステムでは、4月から現在までに実行済みのバッチジョブにて使用したCPUの使用量（以下、CPU使用量とする）と、投入したジョブに設定した実行予定時間より算出されるCPUの予定使用量（以下CPU予定使用量とする）から、各ジョブの実行優先度を決定します。

- 各キュー毎に、実行待ちのジョブにそれぞれ優先順位（63が最優先で0まで）を付け、順位の高いジョブから順に実行します。
- 順位付けの要素は、4月から現在までに実行したバッチジョブのCPU使用量を利用者毎に積算した値と、投入したバッチジョブで要求しているCPU予定使用量を合計したものであり、その値が少ないジョブから順位付けを行います。
- CPU（予定）使用量の算出は以下のとおり行います。CPU使用量についてはシステム（SR8000/MPP, SR11000/J1）別に、利用者毎に積算します。

バッチジョブを実行するキュー		CPU（予定）使用量の算出方法
スカラーキュー（A～F）		ジョブのCPU時間*
拡張記憶（ES）使用	SR8000/MPP	ジョブの実行時間* × 8（CPU）
キュー（A-ES～F-ES）	SR11000/J1	ジョブのCPU時間*
パラレルキュー（P001, P002, P004等）		ジョブの実行時間* × 8（CPU） × ノード数

注：CPU 予定使用量を算出する場合は \* を“実行予定時間”に置き換えてください。実行予定時間はジョブスクリプトに記述した値（#@\$-1T）を使用します。なお、省略した場合はパラレルキューではバッチキューの制限時間、スカラーキュー及び拡張記憶使用キューでは10分が実行予定時間となります。

例：P004 キューで1時間（3,600秒）ジョブを実行した場合のCPU使用量の算出  
 $3,600（秒） \times 8（CPU） \times 4（ノード） = 115,200$

- いずれかのジョブが終了する毎に当該利用者のCPU使用量を積算し、順位の付け替えを行います。大規模計算を行う利用者が不利にならないよう、一定時間毎にCPU使用量に一定の逓減率を掛け、逓減する仕組みとしています。
- キューに並んでいる際の待ち時間は、順位付けの要素に含みません。
- グループコースの専用キューについては、上記の優先順位変更処理は適用されません。但し、実行したジョブのCPU使用量は積算されます。
- ジョブの実行は2つまで同時に可能です。ただし同一のキューでは1つとなります。
- ジョブの優先順位はqstat コマンドの“PRI”の値で確認が可能です。

```
% qstat
  REQUEST      NAME      OWNER      QUEUE  PRI  NICE  CPU  MEM  STATE
99999.mpp-s.cc job.csh  p05xxx     P001   63   20  UNLIM 13824 QUEUED
```