

# TSUBAME 利用クイックレファレンス

(2009/11/04 版)

## 東京工業大学 学術国際情報センター 共用促進事業推進室

TSUBAME の詳細な利用方法については学術国際情報センターWeb ページ

<http://www.gsic.titech.ac.jp/%7Eccwww/index.html>

<http://www.gsic.titech.ac.jp/~ccwww/faq.html>

ならびに「TSUBAME 利用の手引き」を参照してください。

ここでは上記が東工大学内向けに記載されているため、共用促進事業の利用者が利用する場合には、異なっている点等を主に説明します。

### 目次

0.	利用開始前に.....	2
0.1	質問窓口.....	2
0.2	利用可能アプリケーション.....	2
0.2.1	ライセンスが必要なアプリケーション（商用、有償）.....	2
0.2.2	usr/apps/free/ 以下のアプリケーション.....	2
0.2.3	サポート無しアプリケーション.....	2
1.	ログイン環境.....	3
1.1	システムへのログイン.....	3
1.2	ファイルシステムとディレクトリ.....	3
2.	プログラムのコンパイル.....	5
3.	ジョブ管理システムの利用法.....	6
3.1	「qgroup -a」による課金グループの確認.....	6
3.2	「nlge」によるジョブの投入.....	6
3.3	「qstatus -sum」によるバッチジョブの状態の確認.....	7
3.4	ジョブの削除.....	8
3.5	バッチキュー利用不可期間.....	8

## 0. 利用開始前に

### 0.1 質問窓口

質問窓口は [kyoyo@gsic.titech.ac.jp](mailto:kyoyo@gsic.titech.ac.jp) です。

[sodan@cc.titech.ac.jp](mailto:sodan@cc.titech.ac.jp) は学内向けの相談窓口です。利用なさらないでください。

### 0.2 利用可能アプリケーション

#### 0.2.1 ライセンスが必要なアプリケーション（商用、有償）

コンパイラと Gaussian 以外のライセンスが必要なアプリケーション（商用、有償）は利用することができません。ライセンスが必要なアプリケーションは利用者がライセンスを個別にアプリケーションベンダーからライセンスを取得し、各自のホームディレクトリーへインストールすることになります。

#### 0.2.2 /usr/apps/free/ 以下のアプリケーション

/usr/apps/free/に既にインストールされているアプリケーションは共同利用（含む共用促進事業）利用者に自由に使っていただくことができます。

`/usr/apps/free/:`

ACML, FFTW, GAMESS, GROMACS, NWChem, POVray, TINKER, UTChem, MPICH, OpenMPI, JDK, ImageMagick

これらの動作について問題がある場合は [kyoyo@gsic.titech.ac.jp](mailto:kyoyo@gsic.titech.ac.jp) までお問い合わせください。

#### 0.2.3 サポート無しアプリケーション

/usr/apps/nosupport/ および /usr/apps/nosupport10/ に既にインストールされているアプリケーションもご自由に利用ください。しかしながら、当該アプリケーションのご利用に当たってはサポート無しということをご了承ください。

`/usr/apps/nosupport/:`

gnuplot-4.2.3 libxml2-2.6.28 php-4.4.7 php-5.2.2 python2.5.2 R-2.5.0 R-2.5.0\_mpi  
R2.5.0\_vltpi R2.7.2\_vltpi tgc\_formhis

`/usr/apps/nosupport10/:`

abinit\_mp gfarm glpk gnuplot lam libxml2 NVIDIA\_CUDA\_SDK octave oldpgi openmx pgi  
pgi8 php python R ruby sysnoise tgc\_formhis

## 1. ログイン環境

### 1.1 システムへのログイン

SSH を利用して学外向けログインノードへ接続します。

```
ssh ユーザ名@login-gw.cc.titech.ac.jp
```

なお、東工大に訪問されて端末室等で東工大内からアクセスする場合は下記のようにログインしてください。

```
ssh ユーザ名@login.cc.titech.ac.jp
```

ログインノードでの認証に成功すると自動的にインタラクティブノードにログインします。インタラクティブノードでは通常の UNIX の操作、プログラムのコンパイル、ジョブの投入、小規模プログラム（メモリーサイズ 4GB まで、並列数 4 プロセスまで、実行経過時間 30 分まで）の実行（デバッグ、プリポスト処理）が出来ます。

tgg075017 から tgg075032 までがインタラクティブノードです。

合計 16 ノード 256CPU がありますが、MPI 並列の上限は 4 並列、CPU 時間は 30 分です。

5 並列以上の MPI ジョブのデバッグは inno1, inno2, RAM64GB, RAM128GB のバッチキューをご利用ください。

(成果非公開の場合のバッチキューは、pinno1, pinno2, RAM64GB, RAM128GB となります。)

### 1.2 ファイルシステムとディレクトリ

以下のファイルシステムが用意されています。

- ・ /ihome : 共用促進事業専用、バックアップ有り、10Gb Ethernet 接続、ZFS、全ノード共有、強制容量制限無し、100GB 程度に収めてください。
- ・ /iwork : 共用促進事業専用、バックアップ無し、10Gb Ethernet 接続、ZFS、全ノード共有、強制容量制限無し、100GB 程度に収めてください。
- ・ /work, /work2 : 学内共用、バックアップ無し、InfiniBand (10Gbps)、Lustre(並列ファイルシステム)、全ノード共有、強制容量制限 1TB
- ・ /archive : 学内共用、バックアップ無し、InfiniBand (10Gbps)、Lustre(並列ファイルシステム)、全ノード共有、強制容量制限 2TB

/work, /work2, /archive は新規に利用する際には、各自でユーザ名のサブディレクトリを作成してください。  
/work, /work2 ストライピングで高速入出力が出来ますので、利用の手引き p. 25 「3.3.5. Lustre 環境の変更」を参照してストライプのチューニングを行ってください。

```
lfs help setstripe
```

```
lfs help getstripe
```

で設定方法が表示されます。

ホームディレクトリは/ihome/\$LOGNAME です。

プログラムの実行ファイル等はワークディレクトリ/iworkに置くようにしてください。

東工大学内利用者向けと異なり、容量制限がかかっておりません。

ただし、他の課題とディスク全体を共有しておりますので、これから作成する分は、各課題で 100GB までと自粛をお願いします。

100GB を超える容量が必要な課題は、下記の作業領域をご利用ください。ただし利用を希望される方は

kyoyo@gsic.titech.ac.jp まで、必要とする容量を代表者の方がご連絡ください。

tgt074187:/iwork1      15T   32K   15T   1% /iwork1\_01

tgt074188:/iwork2      15T   32K   15T   1% /iwork2\_01

tgt074189:/iwork3      15T   32K   15T   1% /iwork3\_01

tgt074190:/iwork4      15T   32K   15T   1% /iwork4\_01

各課題のグループ（ID の最初の 5 桁）名のディレクトリを上記のいずれかに作成します。

## 2. プログラムのコンパイル

PGI のコンパイラが標準コンパイラとして利用可能です。

この他、インテルコンパイラ、GNU コンパイラ (GCC) が利用可能です。

詳細は利用の手引きおよび以下の Web ページを参照してください。

<http://www.gsic.titech.ac.jp/~ccwww/Announce/compiler.html>

この Web ページでは

PGI

GCC

インテルコンパイラ

MPI 標準環境 (voltaire, MPI1.2 準拠)

MPI2 の利用 (openmpi ベース, PGI 使用)

MPI2 の利用 (openmpi, gfortran 使用)

MPI2 の利用 (openmpi, インテルコンパイラ)

について解説してあります。なお、MPI に関しては可能な限り MPI 標準環境を利用してください。

PGI Fortran でソースコードに STOP 文を含むものをコンパイルして作成した実行オブジェクトファイルの実行後に、標準エラー出力に「FORTRAN STOP」というメッセージが出力されることがありますが、実行およびその結果について影響はありません。また、このメッセージを抑止したい場合は、環境変数 NO\_STOP\_MESSAGE を設定して下さい。なお、値は任意です。

### 3. ジョブ管理システムの利用法

#### 3.1 「qgroup -a」による課金グループの確認

「qgroup -a」コマンドにより課金グループとジョブ実行状態を確認します。GROUP\_NAME の列に表示されるのが課金グループです。次に説明する n1ge コマンドで -g の後に指定するか、環境変数 N1GE\_GROUP に課金グループを指定して下さい。

#### 3.2 「n1ge」によるジョブの投入

n1ge コマンドによりジョブを実行します。-help で詳細な解説が得られます。

どのキューが空いているかは、後述の「qstatus -sum」で確認してください。

利用可能な常設キューは inno1, inno2, RAM64GB, RAM128GB です。

(成果非公開の場合の常設キューは、pinno1, pinno2, RAM64GB, RAM128GB となります。)

この他に事前予約制の hpc1, hpc3 キューが利用可能です。

なお利用可能なキュー名に tes2 がついているものは TESLA を利用可能なノードのみから構成されたキューです。

一例で MPI ジョブを投入する場合のオプションを以下に示します。

```
n1ge -pl 優先度 -noreq -q キュー名 -g 課金グループ -rt 実行上限時間(分) -mpi 全プロセス数:1ノ  
ードのプロセス数 -mem 1プロセス当たりのメモリ(GB) -N ジョブ名 プログラム名
```

```
例) % n1ge -pl 2 -noreq -q inno1 -g 3S090205 -rt 1440 -mpi 16:8 -mem 3.3 -N testjob ./a.out
```

-pl : ジョブ投入時の優先度を指定します。1:通常、2:優先、3:最優先となります。なお、課金係数は未指定時の1, 2, 4倍となります。

-noreq 指定すると指定するとシステム障害等でのジョブ異常終了時の TSUBAME システム側での自動ジョブ再投入を抑制します。

-q キュー名 投入キュー名、共用促進事業では、inno1, (pinno1), inno2, (pinno2), RAM64GB, RAM128GB が利用できます。

inno1, inno2:最大118ノード、2.4GHz CPU, 32GBメモリ

pinno1, pinno2:最大118ノード、2.4GHz CPU, 32GBメモリ (成果非公開用)

RAM64GB:最大16ノード、2.6GHz CPU, 64GBメモリ

RAM128GB:最大2ノード、2.6GHz CPU, 128GBメモリ

-g 課金グループ、qgroup -a で確認ください。

環境変数 N1GE\_GROUP に課金グループを設定すると省略可能

-rt 実行時間上限(分単位) 指定しない場合は、30分になります。

-mpi MPI 並列数:ノード内並列数 プログラム名

-mem プロセスが利用するメモリの最大値(GB単位)、固定小数点指定可能

例) -mpi 4 -mem 3.3 : 4\*3.3=13.2GB を合計で利用

-mem については利用の手引きを良く参照のこと。MPI 並列では1プロセス当たりのメモリ使用量になります。

-N ジョブ名 指定しない場合は自動的にプログラム名に応じたものが付きます。

プログラム名 : 実行モジュールだけでなく、スクリプトも指定可能です。

### 3.3 「qstatus -sum」によるバッチジョブの状態の確認

以下のコマンドでバッチキュー毎の空きノード、空き CPU コア、空きメモリ、利用中ノード、停止ノード、利用不可ノード、合計ノードが出力されます。

```
---
```

QUEUE	FREE NODE	FREE CPU	FREE MEMORY	USED NODE	DOWN NODE	DISABLE NODE	TOTAL NODE
TOTAL	544	3634 CPU	12131 GB	464	3	79	
- bes1	107	845 CPU	2456 GB	9	0	0	116
- bes1tes2	19	76 CPU	397 GB	1	0	0	20
- bes2	74	277 CPU	902 GB	41	0	3	118
- bes2tes2	38	152 CPU	504 GB	27	0	2	67
- cs1	107	214 CPU	2456 GB	9	0	0	116
- cs2	74	148 CPU	902 GB	41	0	3	118
- hpc1	0	0 CPU	0 GB	0	0	67	67
- novice	11	176 CPU	352 GB	0	0	0	11
- RAM128GB	2	32 CPU	256 GB	0	0	0	2
- RAM64GB	15	240 CPU	960 GB	0	0	1	16
- sla1	41	656 CPU	1312 GB	76	0	1	118
- sla1tes2	22	352 CPU	704 GB	37	0	1	60
- sla2	18	288 CPU	576 GB	99	0	1	118
- sla2tes2	9	144 CPU	288 GB	48	0	0	57
- tsubasa	7	34 CPU	66 GB	76	3	0	86

inno1, pinno1, sla1 と inno2, pinno2, sla2 のキューの実体はそれぞれ同一です。

それぞれのキューは両制度で利用可能なキューが異なります。

inno[12] : 共用促進事業、共同利用有償成果公開

pinno[12] : 共同利用有償成果非公開

sla[12] : 東工大学内および共同利用共同研究拠点

なお、qstatus -sum -inno -ram とすると共用促進事業で利用可能なキューに限って出力されます。

```
% qstatus -sum -inno -ram
```

```
---
```

QUEUE	FREE NODE	FREE CPU	FREE MEMORY	USED NODE	DOWN NODE	DISABLE NODE	TOTAL NODE
TOTAL	78	1248 CPU	3168 GB	173	0	3	
- inno1	43	688 CPU	1376 GB	74	0	1	118
- inno2	18	288 CPU	576 GB	99	0	1	118
- RAM128GB	2	32 CPU	256 GB	0	0	0	2
- RAM64GB	15	240 CPU	960 GB	0	0	1	16

### 3.4 ジョブの削除

「qdel\_sla」または「qdelete」コマンドを使ってジョブを削除します。

(例) inno1 で実行しているジョブ (Job-ID 1350) を削除する場合

```
% qdel_sla 1350
```

または

```
% qdelete -c sla 1350
      ^^^          ※キュークラス名
```

または

```
% qdelete -q inno1 1350
      ^^^^^        ※キュー名
```

### 3.5 バッチキュー利用不可期間

平成21年度では、運用が昨年度から一部変更となっております。最大の違いは、東工大内部での利用繁忙期、平成21年12月15日から平成22年2月14日まで、バッチキューにジョブが投入できなくなりました。これを踏まえて、利用計画を修正ください。

利用上不明なことがありましたら、気兼ねなく、kyoyo@gsic.titech.ac.jp まで質問してください。

改訂履歴：

2009/11/04 これまでのテキスト版より、PDF へ変更