

TSUBAME3.0から TSUBAME4.0への移行

Table of contents

1. はじめに	3
1.1. TSUBAME4.0概要	3
2. 利用開始・ログイン方法	4
2.1. アカウントの取得	4
2.2. ログイン方法	4
2.3. ログインノード	4
2.4. TSUBAMEポイントの確認	5
3. 各種環境	6
3.1. ストレージ	6
3.2. 計算ノード	6
3.3. ジョブスケジューラ	7
3.4. ソフトウェア環境	8
4. TSUBAME3.0からのデータ移行	9

1. はじめに

本書は、これまでの TSUBAME3.0 ユーザー向けに、新しい TSUBAME4.0 システムとの違いに重点を置いて説明し、簡単かつすばやく使い始めるられることを目的としています。

1.1. TSUBAME4.0概要

TSUBAME4.0の概要は、[こちら](#)をご覧ください。

2. 利用開始・ログイン方法

2.1. アカウントの取得

本システムを利用するには、予め利用申請を行い、ユーザIDを取得する必要があります。

利用者区分に応じて必要な操作・手続きが異なりますので、詳細は[アカウント取得方法](#)をご参照ください。



利用開始の流れ、アカウント取得方法は TSUBMA3.0 と同じです。

2.2. ログイン方法

ログインノードへのアクセスは、SSHの公開鍵認証方式です。

ログインに使うSSH公開鍵を予めTSUBAME4.0ポータルからアップロードする必要があります。公開鍵の登録の操作は、[TSUBAME4.0ポータル利用の手引き SSH公開鍵の登録](#)を参照ください。



ログイン方法は TSUBMA3.0 と同じです。

2.3. ログインノード

本システムを利用するには、まずログインノードにログインする必要があります。TSUBAME3.0で使用されていたターミナルソフトがそのまま使用できます。

ログイン先には、SSHで接続します。また、ファイル転送はSFTPで接続します。

```
login.t4.gsic.titech.ac.jp
```



ドメインが login.t3.gsic.titech.ac.jp から login.t4.gsic.titech.ac.jp に変更されています。

任意のログインノードにログインしたい場合は、以下のホスト名(FQDN)を指定してください。

```
login1.t4.gsic.titech.ac.jp  
login2.t4.gsic.titech.ac.jp
```



TSUBAME3.0のログインノードは login0, login1 ですが、TSUBAME4.0は login1, login2 です。

2.3.1. ログインノードにおける高負荷プログラムの実行制限について

ログインノード(login, login1, login2)は多数のユーザが同時に利用しているため、CPUを占有するプログラムを実行しないでください。並列計算、長時間な計算は計算ノードを利用してください(qsub/qcrshコマンドを利用する)。

ログインノードでは、1プロセスあたり4GBのメモリ制限を課しています。また、システムに負荷を与えているプログラムはシステム管理者によって予告なく停止させていただきますのでご注意ください。

ログインノードが高負荷で作業しづらい時や、負荷のかかる作業を行うときは、ジョブスケジューラ経由で**インタラクティブジョブ**として実行してください。



ログインノードと計算ノードの使い分けのルールは、TSUBAME3.0と同じです。

2.4. TSUBAMEポイントの確認

コマンドでのTSUBAMEポイントの確認は `t4-user-info group point` コマンドにて確認できます。以下は、TESTGROUPのTSUBAMEポイントを確認する例です。

```
$ t4-user-info group point -g TESTGROUP
gid      group_name      deposit      balance
-----
xxxx    TESTGROUP          5000      800000000
```



TSUBAMEポイントの確認コマンドが `t3-user-info` から `t4-user-info` に変更されています。

3. 各種環境

3.1. ストレージ

TSUBAME4.0 では、ホームディレクトリと2種類のグループディスク 高速ストレージ領域、大容量ストレージ領域 が利用できます。

ホームディレクトリと高速ストレージ領域はSSDの共有ストレージ上に、大容量ストレージ領域はHDDの共有ストレージ上に構築されています。

用途	マウント	容量	ファイルシステム
高速ストレージ領域 Homeディレクトリ	/gs/fs /home	372TB	Lustre
大容量ストレージ領域 共有アプリケーション配備	/gs/bs /apps	44.2PB	Lustre
ローカルスクラッチ領域	/local	各ノード1.62TiB	xfss (SSD)

ローカルスクラッチ領域は各計算ノードのNVMe SSD上にあり、計算途中のテンポラリファイル等に利用できます。



利用可能なローカルスクラッチ領域の容量は、獲得した資源により決まります。
TSUBAME3で利用可能だった共有スクラッチ領域(BeeOND)は廃止されました。詳細については[ストレージ比較](#)をご覧ください。

資源タイプ	ローカルスクラッチ領域 (GiB)
node_f	1660
node_h	830
node_q	415
node_o	200
gpu_1	200
gpu_h	100
cpu_160	830
cpu_80	415
cpu_40	200
cpu_16	83
cpu_8	40
cpu_4	20

3.2. 計算ノード

TSUBAME4.0 の計算ノードは、Zen4 アーキテクチャの第4世代 AMD EPYC 9654 で、1台あたりのコア数が TSUBAME3.0 比で6倍以上と大幅増強されました。

GPU は NVIDIA H100 Tensor Core GPU を4基搭載しています。

	TSUBAME3.0	TSUBAME4.0
演算部名	計算ノード HPE SGI ICE-XA 540台	計算ノード HPE Cray XD665 240台
ノード構成	台あたり	
CPU	Intel Xeon E5-2680 v4 2.4GHz × 2 Socket	AMD EPYC 9654 2.4GHz × 2 Socket
コア数/スレッド	14コア / 28スレッド×2CPU	96コア / 192スレッド×2CPU
メモリ	256GiB	768GiB (DDR5-4800)
GPU	NVIDIA TESLA P100 for NVlink-Optimized Servers ×4	NVIDIA H100 SXM5 94GB HBM2e × 4
SSD	2TB	1.92TB NVMe U.2 SSD
インターコネクト	Intel Omni-Path HFI 100Gbps ×4	InfiniBand NDR200 200Gbps × 4



TSUBAME4.0の計算ノードは、r1n1 r23n11までです。r: 1 23 n: 1 10 or 11

3.3. ジョブスケジューラ

TSUBAME4.0 では、TSUBAME3.0 の UNIVA Grid Engine (UGE) の後継にあたる Altair Grid Engine (AGE) が採用されています。

TSUBAME4.0 の資源タイプは下記の通りです。

資源タイプの種類が増え、各資源タイプで利用できるコア数も増えました。

資源タイプ	使用物理CPUコア数	メモリ (GB)	GPU 数	ローカルクラッチ領域 (GiB)
node_f	192	768	4	1660
node_h	96	384	2	830
node_q	48	192	1	415
node_o	24	96	1/2	200
gpu_1	8	96	1	200
gpu_h	4	48	1/2	100
cpu_160	160	368	0	830
cpu_80	80	184	0	415
cpu_40	40	92	0	200
cpu_16	16	36.8	0	83
cpu_8	8	18.4	0	40
cpu_4	4	9.2	0	20

3.3.1. 定額制ジョブ

TSUBAME4.0では、月単位で計算ノードを準占有利用できる「定額制」を導入しました。

学内利用者および共同利用(学術)のユーザのみが利用できます。

定額制でのジョブ投入は、`-q prior` をつけます。その他のオプションは従量制と同じです。

```
$ qsub -q prior -g [TSUBAMEグループ] スクリプト名
```

オプション	説明
<code>-g</code>	TSUBAMEグループ名を指定します。 スクリプトの中ではなくqsubコマンドのオプションとしてつけてください。
<code>-q prior</code>	定額制ジョブ 実行までに最大1時間の待ちが発生します。

定額制についての詳細は[こちら](#)をご覧ください。



Warning

定額制グループのジョブであっても、`-q prior` の指定がない場合、従量制ジョブとして処理されますので、ご注意ください。

3.4. ソフトウェア環境

3.4.1. 商用アプリケーション

TSUBAME4.0とTSUBAME3.0で利用可能な商用アプリケーションの違いは、[こちら](#)をご覧ください。

TSUBAME4.0では、一部の商用アプリケーションの利用には別途アプリケーション利用料が必要になります。詳細は[利用料の概略](#)のアプリケーション (TSUBAME4.0で一部有償化)をご覧ください。

3.4.2. フリーソフト

TSUBAME4.0とTSUBAME3.0で利用可能なフリーソフトの違いは、[こちら](#)をご覧ください。

3.4.3. TSUBAME3.0で使用していたアプリケーション

TSUBAME4.0とTSUBAME3.0では、コンパイラ、MPI、各種ライブラリが異なるため、そのまま動かすことができません。プログラムをTSUBAME4.0上でコンパイルし直す必要があります。

4. TSUBAME3.0からのデータ移行

TSUBAME3上のデータをTSUBAME4にコピーする例を示します。
以下、コマンドは全て TSUBAME3 上で実行するものになります。
例文中、ux00000 は TSUBAME4 のログイン名に読み替えてください。

- TSUBAME3でSSH鍵ペアを作る

```
$ ssh-keygen -t ecdsa
```

- TSUBAME4ポータルに作成したSSH公開鍵を登録する [登録方法](#)

```
$ cat ~/.ssh/id_ecdsa.pub
```

- TSUBAME3からTSUBAME4にSSH出来ることを確認する

```
$ ssh ux00000@t4-login1  
/home/1/ux00000
```

Note

TSUBAME4のログインノードは2台あり、TSUBAME3からは t4-login1, t4-login2 でアクセスできます。

- TSUBAME3からTSUBAME4ホームディレクトリにファイルを移動する場合

```
$ rsync -az ./dir-to-send ux00000@t4-login1:/home/1/ux00000/dir-to-store
```

- TSUBAME3からTSUBAME4グループディスクへファイルを移動する場合

```
$ rsync -az ./dir-to-send ux00000@t4-login1:/gs/bs/tgz-XXXXX/dir-to-store
```

Tips

TSUBAME4.0のグループディスクは、高速ストレージ領域と容量ストレージ領域があります。詳細は[こちら](#)をご覧ください。

Tips

デフォルト以外のSSH鍵ペアを使用する場合は、下記の通りになります。

```
$ rsync -az -e 'ssh -i /path/to/key' ./dir-to-send ux00000@t4-login1:/home/1/ux00000/dir-to-store
```