

# NVIDIA HPC SDK 利用の手引 き

---

## Table of contents

---

1. はじめに	3
1.1. 利用できるバージョン	3
1.2. マニュアル	3
2. 利用方法	4
2.1. NVIDIA HPC SDKの実行	4
2.2. GPU情報の取得	5

## 1. はじめに

---

### Info

本ページのコマンドライン例では、以下の表記を使用します。

**[login]\$**: ログインノード

**[rNnN]\$**: 計算ノード

**[login/rNnN]\$**: ログインノードまたは計算ノード

**[yourPC]\$**: ログインノードへの接続元環境

本書は、NVIDIA HPC SDK をTSUBAME4.0 で利用する方法について説明しています。また、TSUBAME4.0を利用するにあたっては、[TSUBAME利用の手引き](#)もご覧下さい。サーバの利用環境や注意事項などが詳細に記述されていますので、よく読んでください。

### 1.1. 利用できるバージョン

---

TSUBAME4で利用可能な最新バージョンについてはTSUBAME計算サービスWebサイトの [サポートされているアプリケーション](#) ページをご確認下さい。

研究に支障がない限り、バグ修正の入っている最新版をご利用下さい。

#### 1.1.1. バージョンの切り替え

---

本システムでは、moduleコマンドを使用することでコンパイラやアプリケーション利用環境の切り替えを行うことができます。

例: NVIDIA HPC SDK Version 24.1

```
[rNnN]$ module purge  
[rNnN]$ module load nvhpc/24.1
```

### 1.2. マニュアル

---

- [NVIDIA HPC SDK Version 24.1 \(www.nvidia.com\)](http://www.nvidia.com)
- [CUDA Toolkit Documentation \(www.nvidia.com\)](http://www.nvidia.com)

## 2. 利用方法

### Info

本ページのコマンドライン例では、以下の表記を使用します。

**[login]\$** : ログインノード

**[rNnN]\$** : 計算ノード

**[login/rNnN]\$** : ログインノードまたは計算ノード

**[yourPC]\$** : ログインノードへの接続元環境

### 2.1. NVIDIA HPC SDKの実行

#### 2.1.1. NVIDIA HPC SDK プログラム

GPU非使用時のNVIDIA HPC SDKの使用方法を以下に示します。  
モジュールを利用して、コンパイラの環境、パスを設定します。

```
[rNnN]$ module purge  
[rNnN]$ module load nvhpc
```

NVIDIA HPC SDKのコマンド名、コマンド形式を以下に示します。

コマンド	言語	コマンド形式
nvfortran	Fortran 77/90/95/2003/2008/2018	\$ nvfortran [オプション] source_file
nvc	C	\$ nvc [オプション] source_file
nvc++	C++	\$ nvc++ [オプション] source_file

#### 2.1.2. CUDA・CUDA Fortran

NVIDIA HPC SDKを用いたCUDA・CUDA Fortranの使用方法を以下に示します。  
モジュールを利用して、コンパイラの環境、パスを設定します。

```
[rNnN]$ module purge  
[rNnN]$ module load nvhpc
```

CUDA C、CUDA Fortranのコマンド名、コマンド形式を以下に示します。

CUDA Cのコマンド名とコマンド形式

コマンド	言語	コマンド形式
nvcc	C/C++	\$ nvcc -gencode arch=compute_90,code=sm_90 [options] source_file

CUDA Fortranのコマンド名とコマンド形式

コマンド	言語	コマンド形式
nvfortran	Fortran 77/90/95/2003/2008/2018	\$ nvfortran -cuda -gpu=cc90 [オプション] source_file

#### 2.1.3. OpenACC

NVIDIA HPC SDKを用いたOpenACCの使用方法を以下に示します。モジュールを利用して、コンパイラの環境、パスを設定します。

```
[rNnN]$ module purge
[rNnN]$ module load nvhpc
```

OpenACCのコマンド名、コマンド形式を以下に示します。

OpenACCのコマンド名とコマンド形式

コマンド	言語	コマンド形式
nvfortran	Fortran 77/90/95/2003/2008/2018	\$ nvfortran -acc -gpu=cc90 [オプション] source_file
nvc	C	\$ nvc -acc -gpu=cc90 [オプション] source_file
nvc++	C++	\$ nvc++ -acc -gpu=cc90 [オプション] source_file

OpenACC 利用時の主なオプションを以下に示します。

オプション	説明
-acc	OpenACC指示文に基づきGPUコードを生成します。
-gpu=cc90	ターゲットアーキテクチャを指定します。NVIDIA H100用の実行バイナリを作成します。
-Minfo=accel	OpenACCのコンパイラによる診断情報を出力します。 デフォルトではOpenACCの診断情報は出力されません。

## 2.2. GPU情報の取得

NVIDIA HPC SDKに含まれている `nvaccelinfo` コマンドを用いて、シェアドメモリのサイズ、ウォークサイズ等のGPUの詳細情報を得ることができます。

以下に例を示します。

```
[rNnN]$ module purge
[rNnN]$ module load nvhpc
[rNnN]$ nvaccelinfo

CUDA Driver Version:      12030
NVRM version:             NVIDIA UNIX x86_64 Kernel Module  545.23.08  Mon Nov  6 23:49:37 UTC 2023

Device Number:            0
Device Name:              NVIDIA H100
Device Revision Number:   9.0
Global Memory Size:       99871424512
Number of Multiprocessors: 132
Concurrent Copy and Execution: Yes
Total Constant Memory:    65536
Total Shared Memory per Block: 49152
Registers per Block:      65536
Warp Size:                32
Maximum Threads per Block: 1024
Maximum Block Dimensions: 1024, 1024, 64
Maximum Grid Dimensions:  2147483647 x 65535 x 65535
Maximum Memory Pitch:     2147483647B
Texture Alignment:        512B
Clock Rate:               1980 MHz
Execution Timeout:        No
Integrated Device:        No
Can Map Host Memory:     Yes
Compute Mode:             default
Concurrent Kernels:       Yes
ECC Enabled:              Yes
Memory Clock Rate:        1593 MHz
Memory Bus Width:         6144 bits
L2 Cache Size:            62914560 bytes
Max Threads Per SMP:      2048
Async Engines:            5
Unified Addressing:       Yes
Managed Memory:          Yes
Concurrent Managed Memory: Yes
Preemption Supported:     Yes
Cooperative Launch:       Yes
Cluster Launch:           Yes
Unified Function Pointers: Yes
Default Target:           cc90

Device Number:            1
```

```

Device Name:                NVIDIA H100
Device Revision Number:     9.0
Global Memory Size:         99871424512
Number of Multiprocessors:  132
Concurrent Copy and Execution: Yes
Total Constant Memory:      65536
Total Shared Memory per Block: 49152
Registers per Block:        65536
Warp Size:                  32
Maximum Threads per Block:  1024
Maximum Block Dimensions:    1024, 1024, 64
Maximum Grid Dimensions:     2147483647 x 65535 x 65535
Maximum Memory Pitch:        2147483647B
Texture Alignment:           512B
Clock Rate:                  1980 MHz
Execution Timeout:           No
Integrated Device:           No
Can Map Host Memory:         Yes
Compute Mode:                 default
Concurrent Kernels:          Yes
ECC Enabled:                  Yes
Memory Clock Rate:           1593 MHz
Memory Bus Width:            6144 bits
L2 Cache Size:               62914560 bytes
Max Threads Per SMP:         2048
Async Engines:                5
Unified Addressing:           Yes
Managed Memory:              Yes
Concurrent Managed Memory:    Yes
Preemption Supported:         Yes
Cooperative Launch:           Yes
Cluster Launch:               Yes
Unified Function Pointers:     Yes
Default Target:               cc90

```

```

Device Number:              2
Device Name:                NVIDIA H100
Device Revision Number:     9.0
Global Memory Size:         99871424512
Number of Multiprocessors:  132
Concurrent Copy and Execution: Yes
Total Constant Memory:      65536
Total Shared Memory per Block: 49152
Registers per Block:        65536
Warp Size:                  32
Maximum Threads per Block:  1024
Maximum Block Dimensions:    1024, 1024, 64
Maximum Grid Dimensions:     2147483647 x 65535 x 65535
Maximum Memory Pitch:        2147483647B
Texture Alignment:           512B
Clock Rate:                  1980 MHz
Execution Timeout:           No
Integrated Device:           No
Can Map Host Memory:         Yes
Compute Mode:                 default
Concurrent Kernels:          Yes
ECC Enabled:                  Yes
Memory Clock Rate:           1593 MHz
Memory Bus Width:            6144 bits
L2 Cache Size:               62914560 bytes
Max Threads Per SMP:         2048
Async Engines:                5
Unified Addressing:           Yes
Managed Memory:              Yes
Concurrent Managed Memory:    Yes
Preemption Supported:         Yes
Cooperative Launch:           Yes
Cluster Launch:               Yes
Unified Function Pointers:     Yes
Default Target:               cc90

```

```

Device Number:              3
Device Name:                NVIDIA H100
Device Revision Number:     9.0
Global Memory Size:         99871424512
Number of Multiprocessors:  132
Concurrent Copy and Execution: Yes
Total Constant Memory:      65536
Total Shared Memory per Block: 49152
Registers per Block:        65536
Warp Size:                  32
Maximum Threads per Block:  1024
Maximum Block Dimensions:    1024, 1024, 64
Maximum Grid Dimensions:     2147483647 x 65535 x 65535
Maximum Memory Pitch:        2147483647B
Texture Alignment:           512B
Clock Rate:                  1980 MHz
Execution Timeout:           No
Integrated Device:           No
Can Map Host Memory:         Yes
Compute Mode:                 default
Concurrent Kernels:          Yes
ECC Enabled:                  Yes
Memory Clock Rate:           1593 MHz

```

```
Memory Bus Width:      6144 bits
L2 Cache Size:         62914560 bytes
Max Threads Per SMP:   2048
Async Engines:         5
Unified Addressing:    Yes
Managed Memory:       Yes
Concurrent Managed Memory: Yes
Preemption Supported:  Yes
Cooperative Launch:    Yes
Cluster Launch:        Yes
Unified Function Pointers: Yes
Default Target:        cc90
...
```

ログインノードでnvaccelinfoを実行しても、ログインノードにはGPUが搭載されておきませんので何も表示されません。

qrshやqsubを用いて計算ノードで実行して下さい。

qrsh/qsubに関しましては、[TSUBAME4.0利用の手引き](#)をご参照下さい。