

NVIDIA H100

動作検証レポート



Red Hat、Red Hat Enterprise Linux は、米国およびその他の国における Red Hat, Inc.登録商標または商標です。

NVIDIA は、米国およびその他の国における NVIDIA Corporation の商標または登録商標です。

その他、記載されている会社名、製品名は、各社の登録商標または商標です。

免責条項：本書または本書に記述されている製品や技術に関して、日本電気株式会社またはその関連会社が行う保証は、製品または技術の提供に適用されるライセンス契約で明示的に規定されている保証に限ります。このような契約で明示的に規定された保証を除き、日本電気株式会社およびその関連会社は、製品、技術、または本書に関して、明示または黙示を問わず、いかなる種類の保証も行いません。

目次

NVIDIA H100 動作検証について	3
1 ご利用にあたっての注意事項	3
2 NVIDIA H100 の概要	3
3 検証目的	3
4 動作検証の準備	4
4.1 動作検証システム構成	4
4.2 動作検証済のサーバ構成 (R120j-2M)	4
4.3 NVIDIA H100 増設手順	5
4.4 グラフィクスドライバ適用手順	7
5 検証結果	9
6 関連リンク	9
7 改版履歴	9

NVIDIA H100 動作検証について

1 ご利用にあたっての注意事項

- ・本レポートは動作検証レポートであり、弊社が動作保証するものではありません。
- ・動作確認情報は、各ページに掲載されている評価環境での検証結果に基づいたものです。
- ・導入に際しては個々の環境で十分な確認を実施してください。

2 NVIDIA H100 の概要

NVIDIA H100 をサーバに実装し、NVIDIA vPC ソフトウェアを使用することにより、グラフィックス機能が豊富で、高いパフォーマンスを持った仮想デスクトップインフラストラクチャ(VDI)を実現し、快適なリモートワーク環境を提供することができます。

3 検証目的

今回の検証では、Express5800 シリーズに NVIDIA H100 を搭載し、後述の手配物品、搭載手順、搭載条件を満たした Red Hat Enterprise Linux 8.6 の物理環境下にて、CUDA を使用した GPGPU の基本動作検証を実施した結果を記載します。

4 動作検証の準備

4.1 動作検証システム構成

弊社において検証済みの構成を掲載いたします。なお、下記は一例ですので、お客様の環境や用途にあわせてシステムを構成してください。

4.2 動作検証済のサーバ構成 (R120j-2M)

本章では、動作検証を実施した R120j-2M についてのサーバ手配構成 / 構成に応じた電源ユニットの選択方法 / 動作検証条件について説明します。

なお、NVIDIA H100 は RHEL8 + Express5800/R120j-2M の正式オプション製品となっております。

4.2.1 R120j-2M サーバ手配構成

製品名	対象型名	補足事項
Express5800/R120j-2M	N8100-2988Y	8x 2.5 型ドライブモデル
CPU ボード / 増設 CPU ボード	---	CPU TDP が 300W 以下の CPU のみ搭載可能。 (GPU x1 枚搭載時の場合)
高性能ファン	N8181-209	必須手配品。
高性能 CPU ヒートシンク	N8101-1857	必須手配品。(CPU の手配数と同数必須)
グラフィックスカード電源ケーブル (12+4Pin)	K410-527(00)	必須手配品。NVIDIA H100 に電源供給するために必要。(1 セットで 3 本の補助電源ケーブルが含まれる)
1st ライザカード (3xPCI + 1xGPU 搭載キット)	N8116-112	GPU 3 枚目の搭載に必須
2nd ライザカード (3xPCI + 1xGPU 搭載キット)	N8116-113	2CPU 構成が必須。GPU 1 枚目の搭載に必須
3rd ライザカード (2xPCI)	N8116-115	2CPU 構成が必須。GPU 2 枚目の搭載に必須
電源ユニット(1800W)	N8181-210	必須手配品 (本製品は AC200V 環境のみ使用可能) 2 台手配必須。(冗長化必須)

その他増設オプションについては、Express5800/R120j-2M システム構成ガイドを参照の上、手配ください。

<http://jpn.nec.com/express/systemguide/100guide.html>

4.2.2 NVIDIA H100 搭載時に搭載可能な電源ユニット

搭載するオプション(メモリ、ディスク等)により、システムに要求される電力量が異なります。導入に際しては個々の環境で十分な確認を実施してください。

動作検証条件および搭載制限オプション

分類	GPU 搭載枚数 : 1 枚	GPU 搭載枚数 : 2 枚	GPU 搭載枚数 : 3 枚
CPU	CPU TDP: 300W まで搭載可能	CPU TDP: 225W まで搭載可能	CPU TDP: 165W まで搭載可能
内蔵ドライブ	搭載可能台数 : 8 台以下	搭載可能台数 : 8 台以下	搭載可能台数 : 4 台以下
メモリ	RDIMM : 制限なし	RDIMM : 16 枚まで搭載可能	RDIMM : 8 枚まで搭載可能
増設ドライブ ケージ	搭載不可	搭載不可	搭載不可
PCI カード	制限なし		
防塵フィルタ	搭載不可		
RAID コント	制限なし		
動作環境温度	N8100-2988Y 8x2.5 型ドライブモデル(U.3 NVMe x4/SAS/SATA) : 35 度以下 N8100-2989Y 8x2.5 型ドライブモデル(U.3 NVMe x1/SAS/SATA) : 35 度以下		

補足事項:

- NVIDIA H100 は PCI ライザあたり最大 1 枚まで搭載可能。
- NVIDIA H100 はサーバ 1 台あたり最大 3 枚まで搭載可能。

4.3 NVIDIA H100 増設手順

4.3.1 事前準備

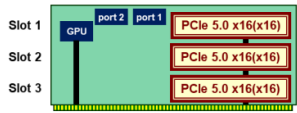
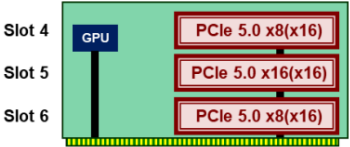
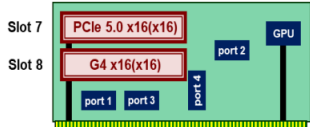
Express5800/R120j-2M への各オプションの実装方法については下記のユーザーズガイドの「2 章 準備」を参照し実装してください。

<https://www.support.nec.co.jp/View.aspx?NoClear=on&id=3170102819>

また、NVIDIA H100 への電源ケーブル接続については、本書 4.3.2 を参照してください。

4.3.2 R120j-2M NVIDIA H100 ケーブル接続 (外部電源)

以下の表を参考に、K410-527(00) グラフィックスカード電源ケーブル(12+4Pin)で NVIDIA H100 を本体装置、およびライザカードの電源コネクタに接続してください。

ケーブル接続先	PCI ライザ概要	PCI ライザ概略図																																
N8116-112 (1st ライザカード)	3xPCI PCI スロット: 3x PCIe 5.0(x16)	 <table border="1" data-bbox="986 766 1433 882"> <thead> <tr> <th>No.</th> <th>Slot</th> <th>PCIe type</th> <th>Description</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>1</td> <td>Full-length/Full-height (FL/FH)</td> <td>PCIe5 x16 (16, 8, 4, 2, 1)</td> </tr> <tr> <td>2</td> <td>2</td> <td>Full-length/Full-height (FL/FH)</td> <td>PCIe5 x16 (16, 8, 4, 2, 1)</td> </tr> <tr> <td>3</td> <td>3</td> <td>Half-length/Full-height (HL/FH)</td> <td>PCIe5 x16 (16, 8, 4, 2, 1)</td> </tr> <tr> <td>4</td> <td></td> <td></td> <td>バックアップ用電源コネクタ (port1)</td> </tr> <tr> <td>5</td> <td></td> <td></td> <td>バックアップ用電源コネクタ (port2)</td> </tr> <tr> <td>6</td> <td></td> <td></td> <td>GPU 電源ケーブルコネクタ</td> </tr> </tbody> </table> <p> ・K410-509(00)が必要 ・NVMe ドライブ CPU 直接接続 16 台構成の場合、搭載不可 </p>	No.	Slot	PCIe type	Description	1	1	Full-length/Full-height (FL/FH)	PCIe5 x16 (16, 8, 4, 2, 1)	2	2	Full-length/Full-height (FL/FH)	PCIe5 x16 (16, 8, 4, 2, 1)	3	3	Half-length/Full-height (HL/FH)	PCIe5 x16 (16, 8, 4, 2, 1)	4			バックアップ用電源コネクタ (port1)	5			バックアップ用電源コネクタ (port2)	6			GPU 電源ケーブルコネクタ				
	No.	Slot	PCIe type	Description																														
1	1	Full-length/Full-height (FL/FH)	PCIe5 x16 (16, 8, 4, 2, 1)																															
2	2	Full-length/Full-height (FL/FH)	PCIe5 x16 (16, 8, 4, 2, 1)																															
3	3	Half-length/Full-height (HL/FH)	PCIe5 x16 (16, 8, 4, 2, 1)																															
4			バックアップ用電源コネクタ (port1)																															
5			バックアップ用電源コネクタ (port2)																															
6			GPU 電源ケーブルコネクタ																															
N8116-113 (2nd ライザカード)	3xPCI PCI スロット: 1x PCIe 5.0(x16) PCI スロット: 2x PCIe 5.0(x8)	 <table border="1" data-bbox="957 1131 1452 1227"> <thead> <tr> <th>No.</th> <th>Slot</th> <th>PCIe type</th> <th>Description</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>4</td> <td>Full-length/Full-height (FL/FH)</td> <td>PCIe5 x8 (8, 4, 2, 1) (x16 コネクタ)</td> </tr> <tr> <td>2</td> <td>5</td> <td>Full-length/Full-height (FL/FH)</td> <td>PCIe5 x16 (16, 8, 4, 2, 1)</td> </tr> <tr> <td>3</td> <td>6</td> <td>Half-length/Full-height (HL/FH)</td> <td>PCIe5 x8 (8, 4, 2, 1) (x16 コネクタ)</td> </tr> <tr> <td>4</td> <td></td> <td></td> <td>GPU 電源ケーブルコネクタ</td> </tr> </tbody> </table> <p>・2CPU 必須</p>	No.	Slot	PCIe type	Description	1	4	Full-length/Full-height (FL/FH)	PCIe5 x8 (8, 4, 2, 1) (x16 コネクタ)	2	5	Full-length/Full-height (FL/FH)	PCIe5 x16 (16, 8, 4, 2, 1)	3	6	Half-length/Full-height (HL/FH)	PCIe5 x8 (8, 4, 2, 1) (x16 コネクタ)	4			GPU 電源ケーブルコネクタ												
	No.	Slot	PCIe type	Description																														
1	4	Full-length/Full-height (FL/FH)	PCIe5 x8 (8, 4, 2, 1) (x16 コネクタ)																															
2	5	Full-length/Full-height (FL/FH)	PCIe5 x16 (16, 8, 4, 2, 1)																															
3	6	Half-length/Full-height (HL/FH)	PCIe5 x8 (8, 4, 2, 1) (x16 コネクタ)																															
4			GPU 電源ケーブルコネクタ																															
N8116-115 (3rd ライザカード)	2xPCI PCI スロット: 1x PCIe 5.0(x16) PCI スロット: 1x PCIe 4.0(x16)	 <table border="1" data-bbox="965 1438 1401 1572"> <thead> <tr> <th>No.</th> <th>Slot</th> <th>PCIe type</th> <th>Description</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>7</td> <td>Full-length/Full-height (FL/FH)</td> <td>PCIe5 x16 (16, 8, 4, 2, 1)</td> </tr> <tr> <td>2</td> <td>8</td> <td>Full-length/Full-height (FL/FH)</td> <td>PCIe4 x16 (16, 8, 4, 2, 1)</td> </tr> <tr> <td>3</td> <td></td> <td></td> <td>バックアップ用電源コネクタ (port1)</td> </tr> <tr> <td>4</td> <td></td> <td></td> <td>バックアップ用電源コネクタ (port2)</td> </tr> <tr> <td>5</td> <td></td> <td></td> <td>バックアップ用電源コネクタ (port3)</td> </tr> <tr> <td>6</td> <td></td> <td></td> <td>バックアップ用電源コネクタ (port4)</td> </tr> <tr> <td>7</td> <td></td> <td></td> <td>GPU 電源ケーブルコネクタ</td> </tr> </tbody> </table> <p> ・2CPU 必須 ・2nd ライザカード必須 ・N8154-175 2x2.5 型 ドライブケース(U.3 NVMe x4/SAS/SATA)を搭載する場合、搭載不可 ・NVMe ドライブ CPU 直接接続 16 台構成の場合、搭載不可 ・3rd ライザカードの Port1-4 と MB の Port 7-8, Port9-10 を接続するためのケーブル付 </p>	No.	Slot	PCIe type	Description	1	7	Full-length/Full-height (FL/FH)	PCIe5 x16 (16, 8, 4, 2, 1)	2	8	Full-length/Full-height (FL/FH)	PCIe4 x16 (16, 8, 4, 2, 1)	3			バックアップ用電源コネクタ (port1)	4			バックアップ用電源コネクタ (port2)	5			バックアップ用電源コネクタ (port3)	6			バックアップ用電源コネクタ (port4)	7			GPU 電源ケーブルコネクタ
	No.	Slot	PCIe type	Description																														
1	7	Full-length/Full-height (FL/FH)	PCIe5 x16 (16, 8, 4, 2, 1)																															
2	8	Full-length/Full-height (FL/FH)	PCIe4 x16 (16, 8, 4, 2, 1)																															
3			バックアップ用電源コネクタ (port1)																															
4			バックアップ用電源コネクタ (port2)																															
5			バックアップ用電源コネクタ (port3)																															
6			バックアップ用電源コネクタ (port4)																															
7			GPU 電源ケーブルコネクタ																															

4.4 グラフィックスドライバ適用手順

4.4.1 事前準備

NVIDIA H100 用の最新のグラフィックスドライバは、下記 Web サイトより入手してください。

<https://www.nvidia.co.jp/Download/index.aspx?lang=jp>

※入手できるグラフィックスドライバのバージョンは時期に依存します。

上記 Web サイトよりグラフィックスドライバを入手する際は下記の点にもご注意の上、適切なグラフィックスドライバを入手してください。

※ご使用になる GPU アクセラレータを[製品タイプ]、[製品シリーズ]、および[製品ファミリー]から正しく選択してください。

異なる GPU アクセラレータのグラフィックスドライバをインストールすると正常に動作しませんので、ご注意ください。

※ご使用になるオペレーティングシステム(OS)を正しく選択してください。

異なる OS のグラフィックスドライバをインストールすると正常に動作しませんので、ご注意ください。

※最新バージョン以外のグラフィックスドライバをご使用になる場合は、下記 Web サイトより入手してください。

<https://www.nvidia.co.jp/Download/Find.aspx?lang=jp>

4.4.2 適用手順

NVIDIA H100 用のグラフィックスドライバは、下記手順を参照の上インストールしてください。

※今回の動作検証では、オペレーティングシステムは[Linux 64-bit RHEL 8]を、CUDA Toolkit は[12.0]を選択し、バージョンが 525.105.17 のグラフィックスドライバをダウンロードして使用しました。ここでは、そのドライバにて確認した手順を記載します。

1. OS を起動後、root 権限のあるユーザーでログインします。

2. nouveau ドライバのロード状態を確認します。下記コマンドを実行してください。

```
# lsmod | grep nouveau
```

nouveau ドライバがロードされていると、グラフィックスドライバのインストールは正常に実行されません。システム起動時に nouveau ドライバが自動的にロードされないよう設定変更を行ってください。

3. elfutils-libelf-devel パッケージがインストールされていることを確認します。

下記コマンドを実行してください。

```
# rpm -q elfutils-libelf-devel
```

パッケージがインストールされていない場合は、インストールしてください。

4. 起動しているカーネルの開発パッケージがインストールされていることを確認します。

下記コマンドを実行してください。

```
# rpm -q kernel-devel-`uname -r`
```

パッケージがインストールされていない場合は、インストールしてください。

5. Systemd ターゲットを multi-user.target に変更します。

6. [4.4.1 事前準備]で入手したグラフィックスドライバを任意のディレクトリに保存します。

7. グラフィックスドライバをインストールします。下記コマンドを実行してください。

```
# sh NVIDIA-Linux-x86_64-525.105.17.run
```

※インストーラの実行中に選択を求めてくる場合がございますので、ご使用の環境に合わせ選択してください。

8. インストール終了後に OS を再起動してください。

5 検証結果

NVIDIA H100 を搭載したサーバにおいて CUDA サンプルによる GPGPU の動作検証をおこなった結果、問題が発生しないことを確認しました。

6 関連リンク

[NEC Express5800/100 シリーズ](#)

[Linux サービスセット \(NEC\)](#)

[NVIDIA 社](#)

7 改版履歴

版数	改版日	変更内容
第 1 版	2023 年 09 月	第 1 版リリース