






技術仕様

Technical Specifications

サーバー向けモデル 仕様一覧

製品名	NVIDIA® Tesla® K80	NVIDIA® Tesla® M40	NVIDIA® Tesla® K40M
外観			
CUDAコア数	4992コア	3072コア	2880コア
プロセッサ周波数	560 MHz (ベース), 562~875 MHz (ブースト)	948 MHz (ベース), 1114 MHz (ブースト)	745 MHz (ベース), 810 / 875 MHz (ブースト)
単精度演算性能	5.6 TFlops (ベース), 8.74 TFlops (ブースト)	7 TFlops	4.29 TFlops (ベース), 5 TFlops (ブースト)
倍精度演算性能	1.87 TFlops (ベース), 2.91 TFlops (ブースト)	-	1.43 TFlops (ベース), 1.66 TFlops (ブースト)
メモリ	24GB GDDR5, 384 bit, 480 GB/sec *1	24GB / 12GB GDDR5, 384 bit, 288 GB/sec *2	12GB GDDR5, 384 bit, 288 GB/sec *3
バス	PCI-Express 3.0 x16	PCI-Express 3.0 x16	PCI-Express 3.0 x16
TDP	300W	250W	235W
補助電源コネクタ	8 pin CPU電源コネクタ ×1	8 pin CPU電源コネクタ ×1	8 pin ×1 / 6 pin ×1
認証規格	EU-RoHS2 / JIG / REACH / WEEE / C-Tick / BSMI / CE / FCC / ICES / KCC / UL, cUL / VCCI		
対応OS	Windows® Server 2012 R2 / Windows® Server 2008 x64 / Windows® 8.1 64bit / Windows® 7 64bit / Linux 64bit (32bit OS はサポート対象外)		
GPU Boost	●	●	●
Hyper-Q	●	●	●
Dynamic Parallelism	●	●	●
ECC 機能サポート	●	●	●
外形寸法 ※ブラケット含まず	267mm x 111mm x 37mm (2スロット占有)	267mm x 111.15mm x 37mm (2スロット占有)	267mm x 111mm x 37mm (2スロット占有)
冷却機構	パッシブ	パッシブ	パッシブ
搭載可能製品	搭載可能製品につきましては、別途お問い合わせください。		
型番	ETSK80-24GER	24GBモデル:ETSM40-24GER 12GBモデル:ETSM40-12GER	ETSK40m-12GER
JANコード	4524076070300	24GBモデル:4524076070423 12GBモデル:4524076070461	4524076070317
保証期間	3年間		

ワークステーション向けモデル 仕様一覧

製品名	NVIDIA® Tesla® K40	NVIDIA® Tesla® K8																
外観																		
CUDAコア数	2880コア	1536コア																
プロセッサ周波数	745 MHz (ベース), 810 / 875 MHz (ブースト)	693 MHz (ベース), 811 MHz (ブースト)																
単精度演算性能	4.29 TFlops (ベース), 5 TFlops (ブースト)	2.13 TFlops																
倍精度演算性能	1.43 TFlops (ベース), 1.66 TFlops (ブースト)	-																
メモリ	12GB GDDR5, 384 bit, 288 GB/sec *3	8GB GDDR5, 256 bit, 160 GB/sec *4																
バス	PCI-Express 3.0 x16	PCI-Express 2.0 x16																
TDP	235W	100W																
補助電源コネクタ	8 pin ×1 / 6 pin ×1	6 pin ×1																
認証規格	EU-RoHS2 / JIG / REACH / WEEE / C-Tick / BSMI / CE / FCC / ICES / KCC / UL, cUL / VCCI																	
対応OS	Windows® Server 2012 R2 / Windows® Server 2008 x64 / Windows® 8.1 64bit / Windows® 7 64bit / Linux 64bit (32bit OS はサポート対象外)																	
GPU Boost	●	●																
Hyper-Q	●	-																
Dynamic Parallelism	●	-																
ECC 機能サポート	●	●																
外形寸法 ※ブラケット含まず	267mm x 111mm x 37mm (2スロット占有)	241mm x 111mm x 18.7mm (1スロット占有)																
冷却機構	アクティブ	アクティブ																
搭載可能製品	<p>下記の必要動作環境を満たした DOS/V PC 及びワークステーション</p> <table border="0"> <tr> <td>CPU</td> <td>CPU Intel® Xeon®/Core i シリーズ、もしくは AMD® Phenom/Opteron 以上</td> <td>必要メモリ</td> <td>4GB 以上のシステムメモリ (8GB 以上のシステムメモリを推奨)</td> </tr> <tr> <td>接続バス</td> <td>PCI-Express x16 に対応した空きスロット 1 つ</td> <td>光学ドライブ</td> <td>ソフトウェア・ドライバインストールのための CD/DVD ドライブ</td> </tr> <tr> <td>必要電源容量</td> <td>最小: 定格出力 600W 以上の出力が可能な電源ユニット 推奨: 定格出力 650W 以上の出力が可能な電源ユニット</td> <td>ハードディスク</td> <td>300MB 以上の空きハードディスク容量</td> </tr> <tr> <td></td> <td></td> <td>その他</td> <td>別途、表示用にグラフィックスボード (オンボードグラフィックス等 *5)</td> </tr> </table>		CPU	CPU Intel® Xeon®/Core i シリーズ、もしくは AMD® Phenom/Opteron 以上	必要メモリ	4GB 以上のシステムメモリ (8GB 以上のシステムメモリを推奨)	接続バス	PCI-Express x16 に対応した空きスロット 1 つ	光学ドライブ	ソフトウェア・ドライバインストールのための CD/DVD ドライブ	必要電源容量	最小: 定格出力 600W 以上の出力が可能な電源ユニット 推奨: 定格出力 650W 以上の出力が可能な電源ユニット	ハードディスク	300MB 以上の空きハードディスク容量			その他	別途、表示用にグラフィックスボード (オンボードグラフィックス等 *5)
CPU	CPU Intel® Xeon®/Core i シリーズ、もしくは AMD® Phenom/Opteron 以上	必要メモリ	4GB 以上のシステムメモリ (8GB 以上のシステムメモリを推奨)															
接続バス	PCI-Express x16 に対応した空きスロット 1 つ	光学ドライブ	ソフトウェア・ドライバインストールのための CD/DVD ドライブ															
必要電源容量	最小: 定格出力 600W 以上の出力が可能な電源ユニット 推奨: 定格出力 650W 以上の出力が可能な電源ユニット	ハードディスク	300MB 以上の空きハードディスク容量															
		その他	別途、表示用にグラフィックスボード (オンボードグラフィックス等 *5)															
型番	ETSK40-12GER	ETSK8-8GER																
JANコード	4524076070225	4524076070249																
保証期間	3年間																	

*1: 1GPUあたりメモリ容量が12GB、メモリバンド幅が240 GB/秒になります。 *2: ECC 機能を有効にすると利用可能なメモリ容量が 6.25%減少し、11.25GB になります。 *3: ECC機能を有効にすると使用可能なメモリは6.25%減少します。 *4: ECC機能を有効にすると使用可能なメモリは10%減少します。 *5: オンボードグラフィックスとの共存ができない場合もございますのでNVIDIA Quadro®/NVS® との組み合わせでのご利用を推奨致します。

お問い合わせ先

株式会社エルザ ジャパン

<http://www.elsa-jp.co.jp>

〒105-0014 東京都港区芝3丁目42番10号 三田UTビル TEL : 03-5765-7391 FAX : 03-5765-7235

● ELSA (エルザ) は、テクノロジー・ジョイント株式会社の登録商標です。● NVIDIA、NVIDIA Tesla は、NVIDIA corporation の商標です。● Windows は米国 Microsoft Corporation の米国及びその他の国における商標及び登録商標です。● その他の商品名は各社の商標または登録商標です。● 仕様などは改良のため、予告なしに変更されることがあります。● 本カタログの掲載内容は2016年4月現在の情報です。

2016.04



NVIDIA® TESLA®

ACCELERATING DATA ANALYTICS, SCIENTIFIC APPLICATIONS,
AND DEEP LEARNING
NVIDIA® TESLA® SERIES

NVIDIA® TESLA® シリーズ 総合カタログ

Design, Built, and Tested by NVIDIA

NVIDIA® TESLA® シリーズ 総合カタログ

Accelerated data centers with Tesla GPUs

NVIDIA® TESLA® K80

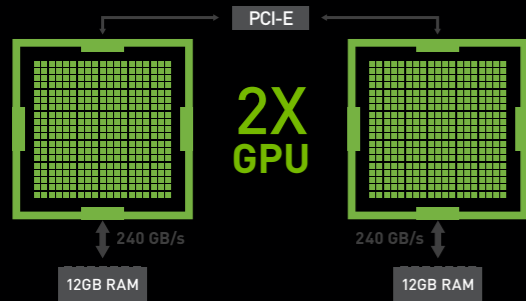
HPC および機械学習用の世界最速 GPU アクセラレータ

NVIDIA® Tesla® K80 は、PCI Express Gen3 に対応し、4992 個の CUDA コアと 24GB の大容量 GDDR5 メモリを搭載し、倍精度演算性能は最大 2.91 TFlops に達します。さらに、GPU Boost 機能を搭載し、エネルギー探索から機械学習まで、かつてない大規模なシミュレーションを高速に実行する GPU コンピューティング環境を実現します。



アプリケーションのパフォーマンスを最大化

2つの GK210 GPU は、高速な PLX スイッチチップで接続され、各 GPU 間のメモリにアクセスする際、この PCI Express スイッチを経由して高速なデータ転送を可能にします。また計 24GB の GDDR5 グラフィックスメモリを搭載し、メモリ帯域は 1GPU あたり 240GB/s、1 ボードあたり 480GB/s に達します。これにより、大規模なデータ処理を必要とするアプリケーションのボトルネックを大幅に軽減します。



GPU Boost

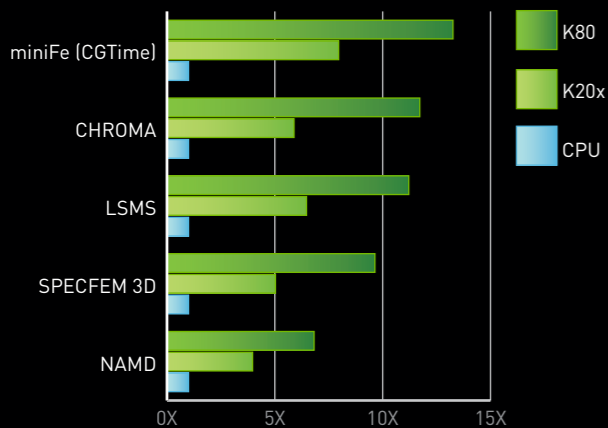
従来定格で動作していたプロセッサ周波数を、プロセッサに余力がある場合にクロックを上げることで、より効率良く GPU を使用します。また、ユーザーはプロセッサがサポートしている範囲でプロセッサ周波数を変更することが可能になり、用途に合わせて、周波数を高めて計算効率を向上させるほか、周波数を抑えて消費電力を抑制することもできます。



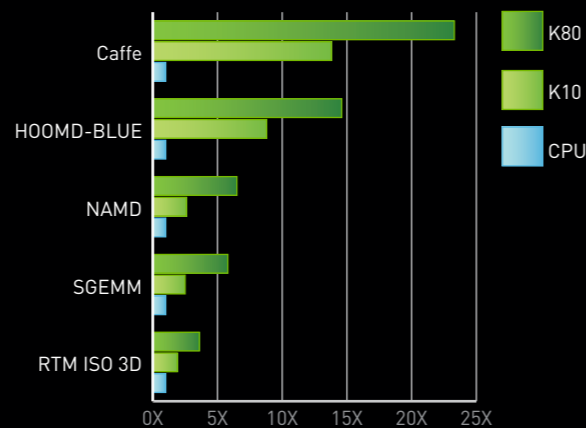
GPU Boost

CPU と比較して 10 倍以上のパフォーマンス

AMBER、HOOMD-BLUE、CAFFE などのアプリケーションにおいて、ハイエンドマルチコア CPU と比べて 10 倍以上の性能向上が達成されています。



CPU: single E5-2697v2 @ 2.70 GHz
GPU: Single K10, K20X or K80, GPU Boost enabled



RTM CPU: dual socket E5-2697v2 @ 2.70 GHz
RTM GPU: Single K10 or K80, GPU Boost enabled

NVIDIA® TESLA® M40

世界最速のディープラーニングトレーニング アクセラレータ

ディープラーニングは、これまで画像認識や自然言語処理に利用されてきましたが、現在は、ニューラルマシン翻訳や画像分類にも利用されています。新興企業から大規模な Web サービスプロバイダーまで、ディープラーニングはエンドユーザーに驚くようなソリューションを提供しています。

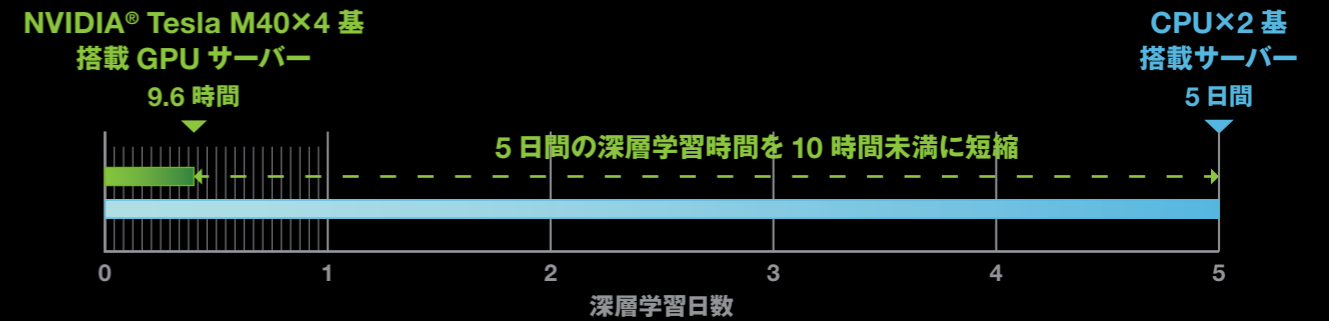
ディープラーニングモデルは、トレーニングの完了までに、数日から数週間もの時間がかかるのが一般的です。このため、サイエンティストは正確性と導入までにかかる時間との間で妥協を強いられています。NVIDIA® Tesla® M40 GPU アクセラレータはディープラーニング トレーニング用の世界最速のアクセラレータであり、トレーニングにかかる時間を劇的に削減するためにデザインされています。

Caffe や Torch といっしょに Tesla® M40 を利用すれば、同じモデルを完成させるのに CPU のみのシステムでは数日かかるのに比べて、たった数時間で実現できます。



CPU 処理に比べ、4 基の Tesla® M40 を用いて深層学習を 13 倍高速化

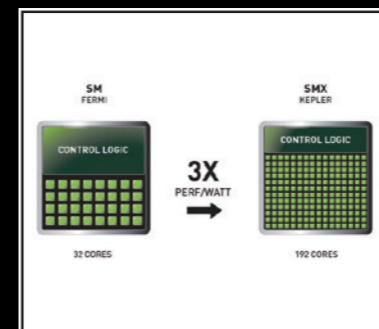
4 基 NVIDIA® の Tesla® M40 を用いることにより、Caffe+AlexNet による深層学習の処理時間が CPU に比べ 9.6 時間まで短縮されています。



Note: Caffe benchmark with AlexNet, training 1.3M images with 90epochs | CPU server uses 2x Xeon E5-2699v3 CPU, 128 GB System Memory, Ubuntu 14.04

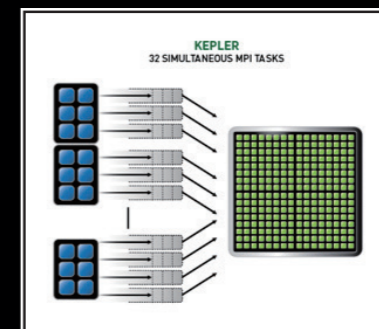
SMX

Kepler アーキテクチャの演算処理ユニットである SMX は、Fermi アーキテクチャに比べ、コントロールロジックを減らしながらプロセッシング・コアを大幅に増加、さらに倍精度演算専用ユニットを追加する事で、従来と同等の消費電力で倍精度 2 倍（単精度 3 倍）の演算性能を実現します。



Hyper-Q

Fermi アーキテクチャ世代の Tesla シリーズでは処理しか実行出来なかった CWD (CUDA Work Distributor) を、Kepler アーキテクチャでは最大 32 のタスクの並列実行が可能となりました。Hyper-Q によって、複数の CPU が一台の GPU 上で同時に作業を開始する事が出来るので、GPU の利用率が劇的に増加し CPU のアイドル時間が大幅に短縮されます。



Dynamic Parallelism

Dynamic Parallelism (動的並列処理) により、GPU が CUDA カーネル実行時に、実行中のカーネル内部から新たなカーネルを生成することが可能になりました。この機能により、ネスティングする複雑な処理や、空間メッシュ分割などの最適化が可能となります。GPU は新たなスレッドを CPU に頼らずに、GPU 自体に動的に生成して、GPU プログラミングを大幅に簡略化し、様々な一般的アルゴリズムをアクセラレートします。

