

24GB の大容量 GDDR5 メモリを搭載
PCIe タイプで GPU サーバー専用のファンレスモデル

NVIDIA® TESLA® P40

人工知能や、インテリジェントマシンの新しい領域である、ディープラーニングは、歴史上、全く新しいコンピューティングモデルの世界を創造しています。革新的な NVIDIA® Pascal アーキテクチャによる GPU パワーは、人工知能の新たな時代の計算エンジンを供給し、ディープラーニングアプリケーションをアクセラレートすることによって、驚くべきユーザーエクスペリエンスを可能にします。

最高の推論処理能力

Tesla® P40 は、ディープラーニング向けに最大のスループットを発揮するように設計されています。Tesla® P40 は、47 TOPS (Tera-Operations Per Second) の推論性能を有し、8 個の Tesla P40 を搭載した 1 台のサーバは、140 個以上の CPU を搭載したサーバの性能に匹敵します。

モデルが正確さや複雑さを増す中で、もはや CPU だけでは、インタラクティブユーザーエクスペリエンスを実現できません。

Tesla® P40 は、最も複雑なモデルであっても、リアルタイムな応答性を、CPU よりも、30 倍以上も低いレイテンシで実現します。



膨大なデータ量に負けない 140 倍の高い処理能力

Tesla® P40 は、最新の Pascal アーキテクチャが採用され、ディープラーニング推論で 47 TOPS 以上もの性能を有します。

8 枚の Tesla® P40 を搭載した 1 台のサーバは、ディープラーニング向けに利用されている 140 個の CPU が搭載されたサーバと同等の性能を有し、結果として、より低いコストでより高いスループットを実現します。

リアルタイム 推論

Tesla® P40 は、最も複雑なディープラーニングモデルを使ったリアルタイムな応答性を、INT8 オペレーションによって、30 倍以上速い推論性能を実現します。


シンプル操作で単一のトレーニングと推論プラットフォームを実現

今日、ディープラーニングモデルは、GPU サーバでトレーニングされていますが、推論は、CPU サーバでデプロイされています。Tesla® P40 は、劇的にシンプル化されたワークフローを提供し、組織が同じサーバで繰り返し、デプロイにも使うことができます。

NVIDIA Deep Learning SDK で 高速なデプロイ

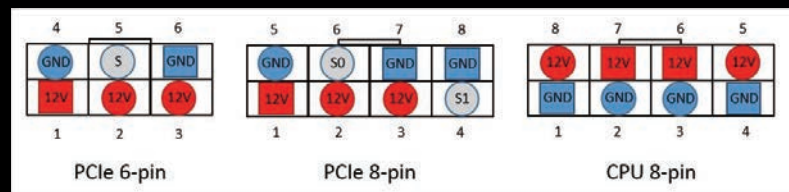
TensorRT は、NVIDIA Deep Learning SDK と Deep Stream SDK からなり、カスタマーがシームレスに新しい INT8 オペレーションやビデオトランスコーディングのような推論能力を利用することをサポートします。

NVIDIA® Tesla® P40 仕様一覧

製品名	NVIDIA® Tesla® P40
外観	
CUDAコア数	3840コア
プロセッサ周波数	1303 MHz (ベース), 1531 MHz (ブースト)
半精度演算性能	47 Tops [Tera Operations Per Second] (最大ブースト)
単精度演算性能	12 TFlops
メモリ	24 GB/バンド帯域幅 347 GB/s)
バス	PCI-Express Gen3 x 16 ^{*2}
最大消費電力	250W
補助電源コネクタ	CPU 8pin コネクタ ×1
認証規格	WHQL / ISO 9241 / EU-RoHS / J-STD / JIG / REACH / HF / WEEE / RCM / BSMI / CE / FCC / ICES / KCC / cUL, UL / VCCI
対応OS ^{*1}	Windows® Server 2016 / Windows® Server 2012 R2 64 / Windows® 10 64bit / Windows® 8.1 64bit / Windows® 7 64bit / Linux 64bit (32bit OS はサポート対象外)
ECC 機能サポート	あり
外形寸法 ※ブラケット含まず	266.7 mm x 111.2 mm x 37 mm (2スロット占有)
冷却機構	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	ETSP40-24GER
JANコード	4524076070539
保証期間	3年間

*1 : 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。

*2 : PCIe 6ピンと PCIe 8ピンと CPU 8ピンコネクタ形状の違い。



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

© 2016 NVIDIA Corporation. All rights reserved. NVIDIA, NVIDIA ロゴ, NVIDIA Tesla, Quadro, NVS, NVIDIA GPU Boost, および CUDA は、NVIDIA Corporation の米国および/または他国における登録商標または商標です。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。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 10 月現在の情報です。