



## NVIDIA TESLA

単一のプラットフォーム。データセンターを無限に高速化。

### コンピューティングの指数関数的な発展

科学的な発見を促進すること、ビッグデータを可視化してインサイトを得ること、顧客に AI ベースのスマートなサービスを提供することが、研究者と技術者にとって日々の課題です。これらの課題を実現するには、さらに複雑で正確なシミュレーションを行い、膨大なデータを処理し、高度なディープラーニングネットワークのトレーニングと運用を行う必要があります。そして、これらのワークロードによって指数関数的に増大するコンピューティングの需要に対応するために、データセンターの高速化も必要です。

NVIDIA® Tesla® は、高速化されたデータセンターのための世界最先端のプラットフォームとして、最大クラスのスーパーコンピューティングセンターおよび企業に導入されています。従来よりも少数で強力なサーバーによって画期的なパフォーマンスが達成されるため、コストを削減しつつ、科学的な発見やインサイトの獲得を迅速化できます。

550 を超える幅広い HPC アプリケーション (上位 15 の HPC アプリケーションすべてを含む)、主要なディープラーニングフレームワークのすべてが GPU 向けに最適化されており、これらを利用する最先端のデータセンターならすべて、Tesla プラットフォームを使用してコストを削減できます。

お客様に最適な NVIDIA データセンター製品を選択してください。

NVIDIA NVLink™ 対応の NVIDIA Tesla V100	NVIDIA Tesla V100 PCIe	NVIDIA Tesla P4	NVIDIA Tesla P40	NVIDIA Tesla P6
				
<p>主な用途 ディープラーニング</p>	<p>主な用途 HPC およびディープラーニング</p>	<p>主な用途 ディープラーニング推論およびビデオトランスコーディング</p>	<p>主な用途 GPU 仮想化 - グラフィックスとコンピューティング</p>	<p>主な用途 GPU 仮想化 - グラフィックスとコンピューティング</p>
<p>演算時間が P100 より最大</p> <p><b>3 倍</b> 高速に</p>	<p>混合ワークロードに対して総所有コスト (TCO) を CPU より最大</p> <p><b>5 倍</b> 削減</p>	<p>推論において CPU より</p> <p><b>40 倍</b> 高いエネルギー効率</p>	<p>1 ボードに最大</p> <p><b>24 基</b> の仮想 GPU</p>	<p>1 ボードに最大</p> <p><b>16 基</b> の仮想 GPU</p>
<p>究極のディープラーニングトレーニングパフォーマンス</p>	<p>混合 HPC ワークロードに対して最大の多用途性</p> <p>メモリを大量に消費する HPC アプリケーション用に 32 GB のメモリ構成</p>	<p>省エネ、薄型でディープラーニング推論のスケールアウト展開に最適</p>	<p>仮想化環境において業界最高のグラフィックス性能</p> <p>仮想化された複数のグラフィックスおよびコンピューティングワークロードを実行</p>	<p>ブレードに最適化されたフォームファクターであらゆる仮想ワークロードで最大のパフォーマンスを発揮</p> <p>フレームバッファを前世代の NVIDIA Maxwell™ の 2 倍に拡張</p>
<p>主な機能</p> <ul style="list-style-type: none"> <li>&gt; 125 TeraFLOPS のディープラーニング用テンソル演算</li> <li>&gt; 15.7 TeraFLOPS の単精度演算能力</li> <li>&gt; 7.8 TeraFLOPS の倍精度演算能力</li> <li>&gt; 300 GB/秒 NVIDIA NVLink™ 相互接続</li> <li>&gt; 900 GB/秒メモリ帯域幅</li> <li>&gt; 32 GB/16 GB HBM2 メモリ</li> </ul>	<p>主な機能</p> <ul style="list-style-type: none"> <li>&gt; 112 TeraFLOPS のディープラーニング用テンソル演算</li> <li>&gt; 14 TeraFLOPS の単精度演算能力</li> <li>&gt; 7 TeraFLOPS の倍精度演算能力</li> <li>&gt; 900 GB/秒メモリ帯域幅</li> <li>&gt; 32 GB/16 GB HBM2 メモリ</li> </ul>	<p>主な機能</p> <ul style="list-style-type: none"> <li>&gt; 22 TeraFLOPS の INT8 推論能力</li> <li>&gt; 5.5 TeraFLOPS の単精度演算能力</li> <li>&gt; 1 デコードおよび 2 エンコード ビデオエンジン</li> <li>&gt; 50 W/75 W 電力</li> <li>&gt; 薄型フォームファクター</li> </ul>	<p>主な機能</p> <ul style="list-style-type: none"> <li>&gt; 24 GB メモリ</li> <li>&gt; 24 の H.264 1080p30 ストリーム</li> <li>&gt; 最大 24 の vGPU インスタンス</li> <li>&gt; PCIe 3.0 デュアル スロット フォームファクター</li> <li>&gt; 250 W 電力</li> </ul>	<p>主な機能</p> <ul style="list-style-type: none"> <li>&gt; 16 GB メモリ</li> <li>&gt; 24 の H.264 1080p30 ストリーム</li> <li>&gt; 最大 16 の vGPU インスタンス</li> <li>&gt; MXM フォームファクター</li> <li>&gt; 90 W (70 W オプション) 電力</li> </ul>
<p>推奨サーバー構成</p> <p>8-ウェイ NVIDIA NVLink ハイブリッド キューブメッシュ (HGX)</p>	<p>推奨サーバー構成</p> <p>1 ノードに 2 ~ 4 基の GPU</p>	<p>推奨サーバー構成</p> <p>1 ノードに 1 ~ 2 基の GPU</p>	<p>推奨サーバー構成</p> <p>1 ノードに 2 ~ 4 基の GPU</p>	<p>推奨サーバー構成</p> <p>1 ノードの GPU 数はブレードサーバーに依存</p>