

ロボティクス、AI、マシンビジョンの性能進化を前進させる産業コンピューティング

ネイサン・ヘップ

GPU、TPU、NVMe、PCIe 6.0といったさまざまな要素が、オートメーションの新しい機会の創出に寄与している。

長年にわたって産業オートメーションは、ロボティクス技術や、データ処理、解析、意思決定のための高度なソフトウェアソリューションの進歩による恩恵を常に受けてきた。ソフトウェアの進化は、さまざまな市場セグメントと応用分野にわたって、さらに革新的な人工知能(AI)ソリューションを生み出している。しかし、その道のりは、平坦でまっすぐなものではなかった。

AIの場合は、リモート、エッジ、モバイル実装に最適ではないフォームファクタやパッケージで提供される、低速なコンピューティング能力が足かせとなって、多くの初期の飛躍的進歩が最近まで実装可能ではなかった。現在、ハードウェア技術の最新の進歩が、判明しているシステム性能能力のギャップを急速に埋める形でこれに交差し、有意義な結果を生成する洗練されたコンピュータビジョンとAIソフトウェアソリューションを、スケーラブルな形で実装するための下準備がようやく整いつつある。

新たな扉を開く CPU、GPU、TPU

コンピューティング能力の進歩について考えるときに、まず頭に思い浮かぶのはCPUとGPUで、それは至極当然のことである。CPUは、コンピューティングプラットフォームの中核要素であり、プログラム命令の実行、オペ

レーティングシステム(OS)の実行、システム入出力(I/O)の管理を司る。標準的なCPUのコア数は、速度とメモリ帯域幅とともに着実に増加しており、それが性能の継続的な向上に貢献している。

加えて、2つの異なる種類のコアを同一デバイスに搭載するCPUが、ますます広く提供されている。それらの高性能コア(Pコア)と高効率コア(Eコア)の利点は、リソースニーズに基づいて負荷が分配されることである。例えば、バックグラウンドの軽負荷のタスクは、より小さく高効率で基本的なタスクに適したEコアに割り当てられる。一方、Pコアは、従来のコアと同様に、性能重視となっている。Pコアは、コンピュータビジョン、ゲーム、機械学習など、演算負荷の高いタスク



の処理に必要なに応じて使用される。

CPUがメインプロセッサであるのに対し、GPU(もともとは、CPUのグラフィカル処理や表示のタスクをオフロードするために開発された)は、グラフィックスと演算を多用する処理に対する定番技術に進化している。最新のGPUデバイスには、複数の処理を並列に実行可能で、多数の小さなタスクを一度に実行するのに最適な、数千ものコアが搭載されている。タスク処理を高速化できるGPUは、AI機能の普及を支える上で役立っている。CPUと同様にGPU技術も、2つの方向に同時に進化している。すなわち、データセン

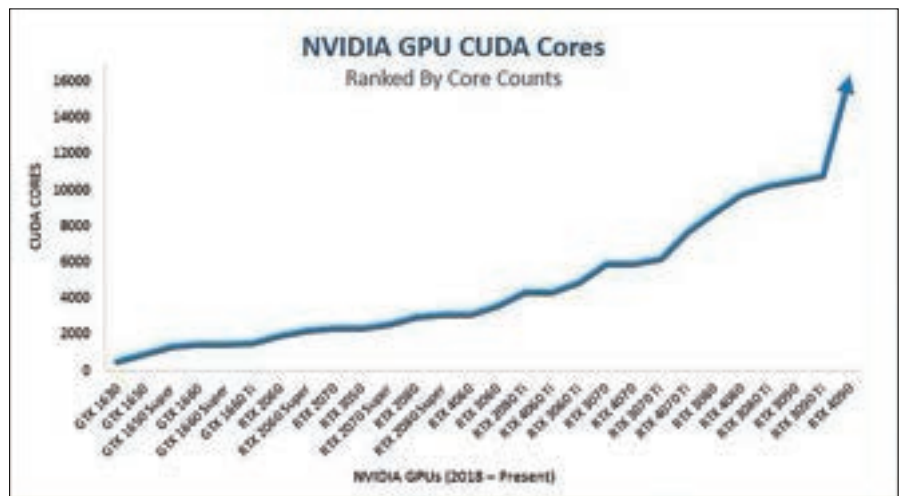


図1 米NVIDIA社のGPU: CUDA コア数順(2018年から現在)

ター、動画編集システム、AIモデルトレーニングの中核要素となる、より大きくて高性能のデバイスと、エッジやモバイル/ロボティクスシステムにおける高性能な処理の実装を可能にする、より小さくて高効率のデバイスである。

これらの2種類のプロセッサに加えて、TPU(Tensor Processing Unit)は、機械学習の活用を目指す分野で、急速に注目を集めている。その名が示すとおり、これらのプロセッサは、ニューラルネットワークの計算を支えるテンソル演算の実行を念頭に、特別に設計されている。TPUは、GPUよりも電力効率が高く、目的に特化して設計されているため、トレーニングと推論のタスクをGPUよりも高速に実行できる。また、TPUは、機械学習フレームワーク「TensorFlow」に組み込まれているため、その機能を活用したいと考える開発者にとってハードルが低い。TPUは、AI分野で使用される産業コンピューティングシステムの性能能力をさらに高める、有望なプロセッサ技術である。そうしたソフトウェア開発の需要に対応するために、急速に成長して成熟していくことが期待されている。

PCIeによるデータ転送の高速化

コンピューティング性能の進歩は、プロセッサ技術だけでは得られない。データが、入力デバイスからシステムへ、また、システム内のデバイスの間を、利用可能なプロセッサの性能を最大限に活用できるだけの速度で流れることが必要である。これができなければ、技術の飛躍的進歩は失われ、その利点は具現化されないままになってしまう。幸い、コンピューティングシステムの他の中核技術も、より大きなエ

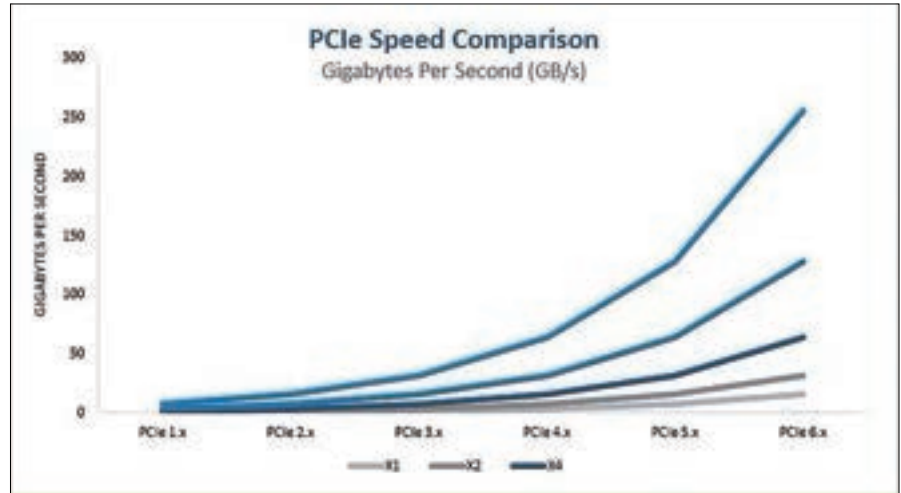


図2 PCIe速度の比較:PCIe 1.xからPCIe 6.xまで(単位:GB/s)

コシステムを支えるために進化している。データベース、メモリ、ストレージの技術が、さらなる高速化、効率化、全体的なシステム性能の向上を先導する役割を果たしている。これらの要素を最先端の処理能力と組み合わせることで、AI、ロボティクス、マシンビジョン、モバイルコンピューティング、エッジコンピューティングの成長促進に必要な環境が構築される。

PCIe (Peripheral Component Interconnect Express)バスは、PC内のほぼすべてのデータのためのインフラ経路である。20年近く前の画期的な誕生から現在に至るまで、多くのコンピューティングプラットフォームの基盤となって、システム内の高速データ転送を実現し続けている。2021年に発表された最新バージョンのPCIe 6.0は、256GB/sの合計双方向転送速度を提供する。これによって、全体的な生産性を最大限に高めるために最新のメモリやプロセッサを利用し続けられるだけの速度で、データを転送することができる。PCIe 6.0は、クラウドコンピューティング、AI、機械学習の分野を大きく後押しする。これらの分野は、データ転送と処理の集中的な負荷を常時システム

に課しつつ、信頼性と堅牢性に優れた相互接続と性能を必要とするためである。より高速なアクセスと転送を可能にして、アイドル時間を最小限に抑えることにより、PCI 6.0は推論とトレーニングの時間も短縮する。

ストレージとメモリの進化

一部の例外を除いて、コンピュータに入力されるデータの大半は、システムメモリまたはRAMに転送される。システムメモリは、データスループットと信頼性の主要なボトルネックになる可能性がある。PCIeバスやその他の周辺インタフェースと同様に、メモリ技術にも、そのような状況を回避するための進歩が継続的に見受けられる。例えば、DDR4 (Double Data Rate 4)メモリの最大データ転送レートは3200MHzで、これは、約25GB/sという、現代のほぼすべての演算タスクに対して十分に高速なピーク転送速度に相当する。

信頼性に関しては、エッジやロボティクスでの使用を意図した多くのコンピューティングプラットフォームは、サイズとフォームファクタが小さい傾向にある。過度のちり、埃、汚染を伴

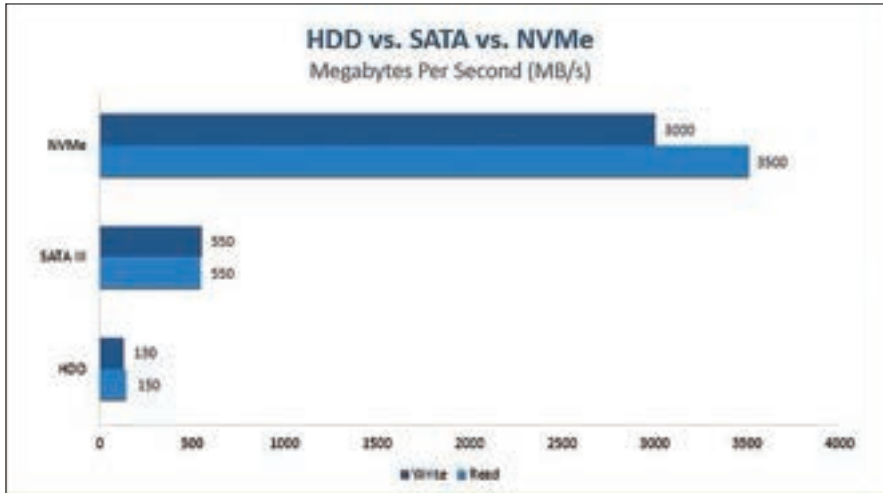


図2 PCIe速度の比較:PCIe 1.xからPCIe 6.xまで(単位:GB/s)

う環境に対応するために、これらのシステムでは、伝導冷却を利用した密閉構造である、ファンレスアーキテクチャの採用がますます増えている。メモリは、従来はこのような用途の一部に対して脆弱だったが、幸いなことに、極端な温度に対応するDDR4メモリが最近登場したことで、この問題も解消されている。-40℃から125℃の温度範囲における安全な動作が保証されれば、人間作業員の安全性を高めるためにロボティクスやモバイルシステムの活用が増加している、過酷な動作環境におけるハイパフォーマンスコンピューティング(HPC)が可能である。

PC性能の最後に残されたボトルネックの1つが、ストレージ周りである。処理のためにデータをストレージメディアから取り出す場合も、処理したデータを保存、記録、将来の使用のためにメディアに書き込む場合も、ハードディスクドライブ(HDD)とソリッドステートドライブ(SSD)の技術は常に、プロセッサ、メモリ、データベースの技術進歩に後れをとってきた。SSDアーキテクチャを対象とするNVMe(Non-Volatile Memory Express)技術の登場によって、ストレージ性能は大

きく向上し、アクセス時間と転送時間はほぼ1ケタ改善されている。これにPCIe 6.0とDDR4メモリ技術を組み合わせ、最新の処理デバイスを補完することにより、今日の高度なコンピュータビジョン、機械学習、AIソフトウェアの要件に十分に答えられる産業コンピューティング性能を、開発者、研究者、ユーザーに提供する、革新的なプラットフォームが完成する。

スモールフォームファクタと高速性

ハードウェアの潜在的性能がソフトウェアの能力に追いついたことで、その恩恵を最も享受するのはどのような分野だろうか。

ロボティクスシステムは、より洗練された用途に日々利用されており、プラットフォームに組み込まれたAI技術によって周囲環境を評価して判断することのできる、適応型ナビゲーションシステムを搭載するロボットを、日常的に目にする。それらのシステムはそのモバイル性により、ほぼ必ずバッテリーで動作するため、最適なコンピューティングを求めつつ、消費電力が重視される。

スモールフォームファクタの組み込み産業用コンピュータは、その要件に応じて、USB 3.0、Thunderbolt、およびPCIeなどの現代的な入力インタフェースを介した高速データ収集に加えて、リアルタイムの処理と意思決定を可能にする。以前は人間が介入しなければ果たせなかったミッションクリティカルでタイムセンシティブな用途に、今ではロボティクス機能の支援が追加されている。

シミュレーションとデジタルツインの用途では、稼働中の製造ラインから取得した実データを使用して、製造およびファクトリオートメーション環境を複製し、AIモデルのさらなるトレーニングや、ソフトウェア変更や新しい装置またはデバイスを追加する影響の正確な評価を、実稼働の完全性を損なうことなく行うことができる。これを実現するのは、今日のコンピューティング技術が可能にした高速データ転送であり、その結果として、最終的には生産性、安全性、製品品質が向上し、誰もが恩恵を受ける。

オートメーション市場範囲の拡大

マシンビジョンの世界では、スループットと精度が、製造、検査、組み立ての処理における最重要項目である。より高度なアルゴリズムとソフトウェア機能を活用する能力が、オートメーションの新しい扉を開き、それが利用される市場の範囲を拡大し続けるだろう。その飽くなき探求に支えられて、ソフトウェアはハードウェアの限界を押し上げ、未来とその先の処理ニーズに応えるべく、将来世代におけるそのさらなる進化を推進し続けるに違いない。

著者紹介

ネイサン・ヘップ(Nathan Hepp)は、米CoastIPC社のクリエイティブディレクター。