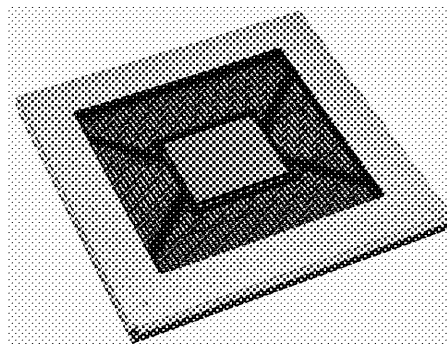


AI半導体開発加速



第2世代の半導体「MN Core2」

人工知能(AI)モデルを学習、推論するのに必要な半導体。現在そうした用途の半導体は米エヌビディアの画像処理半導体(GPU)の独壇場だ。一方、AIの活用が広がる中、独自にAI半導体を開発する企業が増えている。エヌビディアとは違う競争軸を打ち出し、AI市場の拡大を捉える。

(小林健人)

PFN、深層学習に特化

現在AI学習で使われ、AI用途に最適化されるGPUなどは、汎用性が高い。そのため、汎用性を担保した上で、大規模言語モデル(LLM)の登場によりAIの計算能力や電力効率を高める。ただ、GPUはIモデルが大型化する

のに伴いユーザーが求める計算能力は高まり、消費電力が増加傾向にあるが、GPUの消費電力は、プリファードネットワークが設計した第2世代の半導体「MN Core2」

GPUより省電力

力の効率化が追いついていないのが現状だ。元々、半導体はエネルギー多消費産業と言われているが、このままAI開発が進めばさらに多くの電力を使用するという課題を抱えている。そこでプリファードネットワークス(PFN、東京都千代田区)はソフトウェアの開発力を生かし、深層学習に最適化した半導体を開発する。ネットワーク制御などの機能をソ



フトで置き換え、従来に比べて多くの演算機能を搭載した。用途を深層学習に特化することで消費電力を抑えた。

24年中には、第2世代の自社設計半導体を組み込んだサーバーをクラウドサービスとして他社に提供する方針だ。自社で開発するAIソフトもサーバー上で動かせるようにし、ソフトウェア開発キット(SDK)も公開していく(小倉崇浩MN Core事業経営企画室担当VP)。また同社はインター

機能で差別化、エヌビディア追う

ネットインシアティブ計を重視する。端末側(IIJ)や北陸先端科学技術大学院大学と、必要な機能を備えたソフトウェア半導体を開発し、26年以降に最適化し、製造24年からは数十億規模にも製品化する。韓国サムスン電子のAIモデルに対応し、25年にも始める2ナ「SAKURA2」(ナノは10億分の1)の量産を始める。サキのファウンドリー「ヤシノガ・ダスグプタ」を活用。PFNの最高経営責任者(CEO)は「引き合いは強」(野澤一朗VPコンピ)と話す。担当最高技術責任者「AIの市場は今後も(CTO)は「同世代の拡大が予想される。GPUより電力当たり情報技術産業協会の性能と価格面で上回(JEITA)によれば、生成AI市場の世界需要額は30年には2110億ドル(約30兆円)に達し、23年の約20倍になるという。市場拡大に合わせ、計算のAI処理に特化した能力と消費電力のバランスを探る動きも出てソフトを中心とした設計をうたう。