

ビッグデータ/AI時代は計算力! 画像検索サーバ65倍速!

松村 秀敏

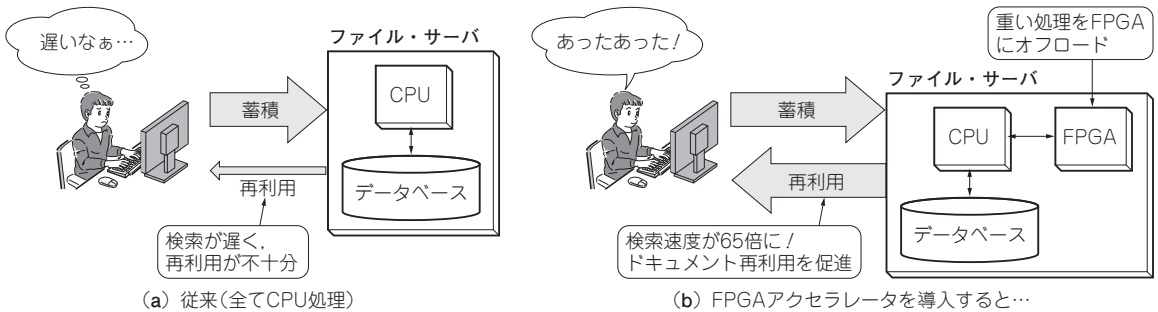


図1 ビッグデータ/AI時代のサーバは処理性能が重要…用途に応じた専用計算回路をFPGAで実現すれば劇的に性能を上げられることがある

人工知能やビッグデータ処理など、大規模データを高速処理するニーズが増大し続けています。一方、ムーアの法則に従った単体CPUコアの性能向上は限界を迎つつあります。このような状況の中、サーバ処理の高速化手段としてFPGAが近年、注目を集めており、さまざまな試みが行われています。

最も有名などころでは、マイクロソフトの事例でしょう。マイクロソフトでは、同社のデータ・センタにFPGAアクセラレータを導入し、ウェブ検索エンジンBingの高速化を試みました。結果として、同程度の性能を実現するためのサーバ台数を半減し、データ・センタのコストを大幅に削減可能であることを示しました⁽¹⁾。

単に高性能を実現するだけならば、大量のサーバで並列に処理したり、GPUを使うといった手段も考え

られます。しかし、FPGAを使ってターゲット・アプリケーションに特化した専用回路を設計することで、CPU/GPUより高い電力効率を実現できます。アプリケーションによっては、性能面でもCPU/GPUを上回ることがあります。

従来、FPGAを使うには、対象アルゴリズムをVerilog HDLなどのハードウェア記述言語(HDL: Hardware Description Language)で記述する必要があり、ハードルが高いという問題がありました。しかし近年の高位合成技術の進歩や、OpenCLとはじめとする開発環境の整備によって、ハードウェア・エンジニア以外のエンジニアにとってハードルが低くなってきています。これも、FPGAのサーバ処理の高速化への進出を後押しする要因となっています。

本稿では、FPGAによるサーバ処理の高速化の事例を紹介します(図1)。

システム構成

本稿で解説する文書検索システムの構成を図2に示します。FPGAアクセラレータによって画像検索が65倍以上高速化し、効率的な文書検索システムを構築できました。

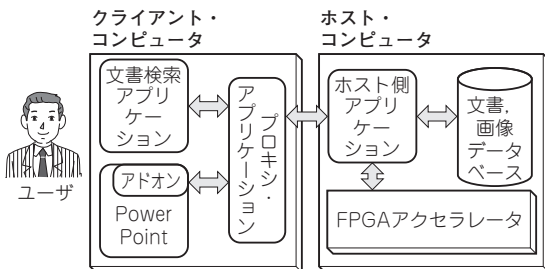


図2 画像検索が65倍以上高速化! FPGAアクセラレータ搭載文書検索システムの構成

● FPGA ボード

今回は、Bittware社のS5-PCIe-HQというFPGAボー