

演算コア数千個! 画像処理もそれ以外も

NVIDIA製

# 第2章 最新GPUのメカニズム

森野 慎也

表1 コア数は数千個! 最近のGPUボード  
すべてNVIDIA製で、汎用演算をこなすしくみ「CUDA」に対応する

デバイス名	GeForce GTX780	Tesla K40c	Quadro K6000	GeForce GTX980	GeForce TITAN-X	Quadro M6000
アーキテクチャ	Kepler			Maxwell		
CUDAコア数	2304	2880		2048	3072	3072
メモリ・バンド幅[バイト/s]	288G			224G	336G	317G
メモリ・サイズ[バイト]	3G	12G		4G	12G	
消費電力[W]	250	235	225	165	250	

並列プロセッサのかたまりであるGPU (Graphics Processing Unit) は、画像処理が得意です。応用として、マシン・ビジョン(外観検査装置)、医療、加えて、車の自動運転やロボットの制御に用いられるコンピュータ・ビジョンなどで、GPUが注目されています。

GPUは並列処理を高速に行え、CPUと比較して数倍の演算性能と、広帯域のメモリ・バンド幅を持ちます。その性能を生かして、画像処理だけでなく、汎用処理を行う並列プロセッサGPGPU (General Purpose Computation on Graphics Processing Unit) としても使われつつあります。

GPUで汎用処理を行う環境も整ってきています。例えばCUDA (Compute Unified Device Architecture) と呼ばれるNVIDIAのGPU専用のプログラミング環境があります。これを使うと、C/C++、Fortranなどで、GPUで効率よく並列処理ができます。

本稿では、GPU (NVIDIA製) の特徴、アーキテクチャと、CUDAのしくみを解説します。(編集部)

## GPUの特徴

### ● CPUの数倍の演算性能&メモリ・バンド幅が超広い

GPUは、CPUと比較して数倍の高い演算性能、および、超広帯域のメモリ・バンド幅を持っており、高い処理性能を実現できます。

GPU上での処理の並列度が非常に高いことが、特

徴の一つに挙げられます。最近のGPUは、一つのデバイス上に、数百個から数千個のコアを搭載しています。表1に最近のGPUボードを示します。

### ● 並列計算はお手のもの

GPUはしばらくグラフィックス専用用途で使われてきましたが、並列処理が得意な点を生かして、GPUを汎用の並列プロセッサとして使うようになってきました。そのために使うしくみがCUDAです。表1に示したGPUはすべてCUDAに対応しています。

### ● GPUで汎用計算を行うためのしくみ「CUDA」

CUDAは、NVIDIAのGPU上で汎用演算を行うためのしくみです。“Compute Unified Device Architecture”の略です。グラフィックスではなく、汎用演算(Compute)を行うためのアーキテクチャです。転じて、NVIDIAのGPUで動作する並列処理プラットフォームと説明しています。

CUDAを使うには、CやC++などでプログラムを作ります。CUDA用に命令が拡張されているので、例えば、CUDAを利用できるC/C++はCUDA C/C++と呼ばれ、CUDAを利用できるFortranは、CUDA Fortranと呼ばれます。GPUをGPGPUとして使いはじめた当初、OpenGL、DirectXなどのグラフィックス・ライブラリの機能を汎用処理に転用していたのは大きく異なり、とても使いやすくなっています。

CUDAに対応したNVIDIAのGPUは、初期のTesla