ご購入はこちら

第2章 ハードウェア方式のSIMD/SIMTの違いや，ソフトウェア環境の思想の違いを見てみる

GPUのアーキテクチャとプログラミング・モデル

塩谷 亮太

表1　GPUの種類とそれぞれのプログラミング・モデル
各GPUはハードウェアとソフトウェアからの見え方が異なる場合があり，厳密にはこの表に当てはまらない場合がある

GPUの種類	ハードウェアの方式	プログラミング・モデル	ソフトウェア環境
NVIDIA (Volta)	SIMT	SPMD	CUDA/OpenCL
AMD (GCN，RDNA)	SIMD	SIMD命令	OpenCL
インテル (GEN)	SIMD	SIMD命令	C for Metal/OpenCL

● 本稿で解説すること

NVIDIAやAMD，インテルなどから発売されている代表的なGPUは，SIMDやSIMTと呼ばれるハードウェアの方式を採用しています．本稿では，このSIMDやSIMTとは一体どのような方式であるのかを説明し，そこにはどのような意味があるのかを説明します．また，各社のGPUには，CUDAやOpenCL，C for Metalなどのソフトウェア環境が存在します(**表1**)．これらはそれぞれプログラミング・モデルも異なっています．

この各環境におけるプログラミング・モデルは，GPUのハードウェアの方式と深く関係があります．本稿ではそのようなプログラミング・モデルとハードウェアの方式の関係や，それらの違いについて説明していきます．

● ビギナでも読めるように配慮した

本稿では，前提としてCPUによるプログラム実行の基本的な仕組みをある程度知っていることを想定しています．例えば，命令，PC(プログラム・カウンタ)，レジスタといった単語の意味については解説しません．

しかし，普段からハードウェアにあまりなじみがなく，これらの言葉を知らない場合であっても，次のように言葉を読み替えていただければ，おおよそは理解できるのではないかと思います．

- 「命令」を「プログラムの各行に書かれている処理」に
- 「PC」を「プログラムのいま実行している行の番号(を記憶している変数)」に

例えば，

CPUやGPUは，メモリからPCが指すアドレスの命令を読み出し，読み出した命令を解釈して演算する

という文は，

CPUやGPUは，プログラムから実行対象となる行に書かれている処理を1つ読み出し，読み出した処理を解釈して演算する

と読み替えてもおおよそ問題ありません．上記のような読み替えを行いながら本章を読んでいただければ，おおよその意味はつかめるのではないかと期待しています(やや無理があるとは思いますが，雰囲気だけでも通じればと願っています)．

ハードウェア構成のキモ…SIMDとは

現代の主なGPUはSIMDやSIMTと呼ばれているハードウェア構成を採用しています．この中でもSIMTはSIMDに基づいた概念であるため，以下ではまずSIMDから説明を始めます．

SIMTについては，そのプログラミング・モデルと関係が深いため，後半で説明します．

SIMDという言葉は「Single Instruction stream/Multiple Data stream」を略したものです．これはコンピュータ・ハードウェアの構成方式の1つです．この分類は，もともとフリン(M. J. Flynn)氏の論文に由来します[1]．

● フリン氏はプログラムを処理する際の命令やデータの流れに応じて4つに分類した

フリン氏はプログラムを処理する際の命令の流れ(stream)とデータの流れに着目し，それらが同時に1つあるか複数あるかに応じてコンピュータの方式を4