

マルチ GPU に対する理論モデル及び GPU アルゴリズムの解析と実現

小山田 徹[†] 和田 幸一^{††} 藤本 典幸^{†††}

[†]法政大学大学院理工学研究科応用情報工学専攻 ^{††}法政大学理工学部応用情報工学科

^{†††}大阪府立大学大学院工学研究科電気・情報系専攻

1. はじめに

グラフィックの画像処理を行うコンピュータの部品に GPU(Graphics Processing Unit)があり, GPU は強力な計算能力を持っている. 本研究では GPU を複数搭載したマルチ GPU の理論モデルを構築し, マルチ GPU 上で動作するアルゴリズムの解析と実装を行う.

2. GPU の特徴

GPU にはグローバルメモリとシェアードメモリが存在する, グローバルメモリは容量は大きくレイテンシが大きい. シェアードメモリは容量は小さくレイテンシも小さい. GPU のプログラミングにはこの 2 つのメモリを効率的に使用することが重要になる. またコンピュータ 1 台に複数の GPU を搭載し, 並列に動作させる技術をマルチ GPU という. 現在扱っているマルチ GPU では GPU を 3 枚搭載し, GPU から GPU へメモリの転送を行うことができるようになっている.

3. 理論モデル

今回で提案するマルチ GPU に対する理論モデルは, 以前研究室で我々が提案したシングル GPU に対する理論モデル[1]を発展させたものである. 本研究で提案するモデルは, GPU の MP(マルチプロセッサ)の数や MP 内で同時に実行できるスレッドの数, グローバルメモリとシェアードメモリの容量やレイテンシなどを考慮している. またコアキャッシング, バンクコンフリクトなどのメモリアクセスの特徴も捉えている. それに加えて GPU の枚数や CPU-GPU 間, GPU-GPU 間のメモリ転送にかかる時間を考慮した構成になっている. ま

た各 GPU のグローバルメモリに入力を分割して配置できるため, 評価するアルゴリズムによって初期入力を変えることができる. マルチ GPU に対する理論モデルを図 1 に示す.

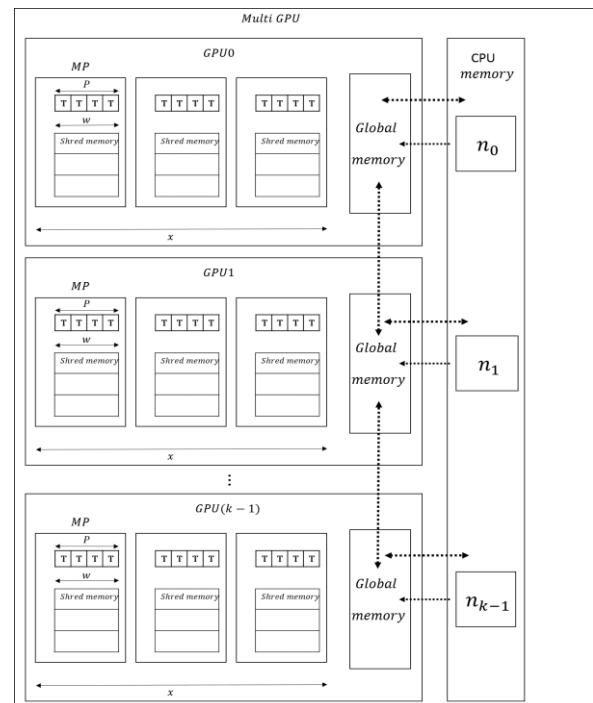


図 1: マルチ GPU に対する理論モデル

4. 結果

マルチ GPU 上で動作するアルゴリズムを評価するための理論モデル(図 1)を提案した. また畳み込み計算のアルゴリズムをこの理論モデルを用いて評価した. そして畳み込み計算アルゴリズムをマルチ GPU 上で実装し, 計算時間を計測した.

参考文献

[1]: 鈴木, 遠藤, 和田: GPGPU に対する理論モデル, 2013 電子情報通信学会総合大会, ISS 特別企画, 学生ポスターセッション, DK-1 (2013-03)