

巻頭言（その2）：第二軍の大輸送

中島 研吾

東京大学情報基盤センター

Oakforest-PACS の後継機 (OFP-II) は、目標ピーク性能を 100-200 PFLOPS として検討を進めて来たが、現実的な消費電力・エネルギー（つまり電気料金）でこれを実現するには、「GPU 等の演算加速器を搭載した計算ノードを主とする」ことが必須であり、既に 2021 年秋頃にはこの方針に決していた。その後、利用者の皆様に向けては、加速器搭載システムへのアプリケーション移行に関するアンケートも実施した。当時、既に「脱炭素化」という世界的潮流があったが、2022 年になってからの電力単価高騰もあり、結果的には正しい選択であった。なお、OFP-II では汎用 CPU のみを搭載した計算ノードも一定数導入する予定である。

TOP500¹⁰のための HPL (High Performance Linpack) ¹¹性能測定時の消費電力あたり演算性能 (GFLOPS/W) によるランキングが Green500¹²である。Wisteria/BDEC-01 のうち、Odyssey (Fujitsu/Arm A64FX) は 15 GFLOPS/W, Aquarius (Intel Icelake+NVIDIA A100) は 24 GFLOPS/W で、Aquarius の効率が約 60%良いことになる。2022 年 11 月現在の Green500 のトップのシステム (Henri) は NVIDIA の最新の GPU (H100) を搭載しており、消費電力あたり性能は 65 GFLOPS/W, Odyssey の 4 倍以上である。

GPU への移行に当たって問題となるのが、アプリケーションの移植である。OFP の利用者は約 3,000 人おり、全てのアプリケーションの移植を完了するには最低でも 18 ヶ月を要すると見込まれる。OFP-II の運用が 2024 年 4 月に開始するとして、そこから逆算すると 2022 年秋には移植のための活動を始めなければならず、そのためにはそれ以前に OFP-II で採用する GPU を決定しておく必要があった。そこで 2022 年年明け早々から GPU 選定のためのベンチマークと附属書類の整備を開始し、表 1 に示すような 7 種類のベンチマークを実施することになった。いずれも計算科学系アプリケーションである。OFP-II では、Wisteria/BDEC-01 で実施している「計算・データ・学習」融合という路線は継続するものの、ワークロードの中心は「計算」であり、GPU 選定にあたって最もその点を最も重視した。

表 1 GPU 選定のためのベンチマーク概要

Name of the Code	Description	Lang.	Parallelization	GPU	Category
P3D	3-D Poisson's Equation by Finite Volume Method	C	OpenMP		A
GeoFEM/ICCG	Finite Element Method	Fortran	OpenMP, MPI		
H-Matrix	Hierarchical-Matrix calculation	Fortran	OpenMP, MPI		
QCD	Quantum-Chromo Dynamics simulation	Fortran	OpenMP, MPI	CUDA	B
N-Body	N-Body simulation using FDPS	C++	OpenMP, MPI	CUDA	
GROMACS	Molecular Dynamics simulation	C++	OpenMP, MPI	CUDA, HIP, SYCL	
SALMON	Ab-initio quantum-mechanical simulator for optics and nanoscience	Fortran	OpenMP, MPI	(OpenACC)	A

¹⁰ <https://www.top500.org/>

¹¹ <https://www.top500.org/project/linpack/>

¹² <https://www.top500.org/lists/green500/>

右端の「Category」という列は、A：汎用 CPU 向けベンチマーク、B：既に GPU 化されたもの、という分類である。10 年ほど前までは、GPU プログラミングと言えば CUDA が必須で、これが大きな障壁となっていたが、その後、指示行ベースの OpenACC、OpenMP、また最近は標準言語（Standard Language）も整備され、プログラミングも容易になっている。絶対的な計算性能が高いことはもちろんであるが、「A Category」では、標準言語等による最低限の変更によるコードと OpenACC、CUDA によるフル最適化されたコードの性能比が 1 に近いことも評価基準の一つとした。

2022 年 2 月に各社によるベンチマーク開始、5 月に結果提出、ヒアリング・審査の後、6 月に NVIDIA の GPU（H100 またはその後継機種）を選定することに決定した。選定のポイントとなったのは：

- ① ベンチマークの演算性能
- ② 移植（ポータリング）のしやすさ
- ③ サポート体制
- ④ Fortran への対応

の 4 点である。科学技術計算においては、我が国では多くのプログラムが Fortran で記述されているため、④は特に重要であった。

GPU への移行に当たっては、各利用者がアプリケーションの移植を実施することを原則としている。そのため、ミニキャンプ、オンライン講習会を以前より頻繁に実施する他、「相談会」を毎月実施し、JCAHPC・NVIDIA 関係者が対応に当たっている。ミニキャンプや相談会は、共通の技術的課題を抱えている他の利用者とのネットワーク形成にも大きく貢献している。また、ミニキャンプでは、JCAHPC の OB・OG を含む他大学教員、関連会社社員の皆様にもメンターとして加わっていただいている。この場を借りて感謝の意を表する次第である。移行のためのポータルサイト¹³も整備されており、講習会の資料、ビデオ動画を初め、様々な情報が集約されている。是非一回覗いてみていただきたい。環境としては、現在は Wisteria/BDEC-01（Aquarius）を使用しているが、最新の H100 を搭載したシステムを利用できるよう調整を進めている。

OFPI, Odyssey ユーザーのプログラムはほとんどが OpenMP+MPI によって並列化されているため、利用者の皆様には、GPU 移行にあたっては OpenACC を利用することをお勧めしている。また標準言語は、特殊な GPU プログラミングの知識は不要であり、容易に GPU 化可能である。図 1 は心臓電気生理学（cardiac electrophysiology）の有限体積法シミュレーションを

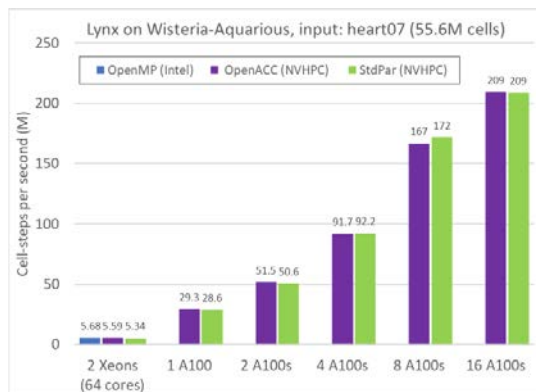


図 1 心臓電気生理学有限体積法シミュレーションの計算性能（時間当たり処理量）、OpenACC と標準言語（StdPar）の比較、Wisteria/BDEC-01（Aquarius）使用 [1]

¹³ https://jcahpc.github.io/gpu_porting/

Wisteria/BDEC-01 (Aquarius) で実施した結果 [1] で、計算時間の大部分は疎行列を係数行列とする連立一次方程式の前処理付き共役勾配法による求解である。図は、時間当たりの処理量を示している。Aquarius は 1 ノード 8GPU (NVIDIA A100) を搭載しており、最大 2 ノードを使用している。1GPU の性能は、CPU1 ノード (Intel Icelake) の約 5 倍、また OpenACC と標準言語 (StdPar) の計算性能はほとんど変わらないことがわかる。比較的単純な演算では、標準言語でも十分高い性能を得られる。

この他、OFP, Odyssey の大口ユーザーのグループ、広く使用されているコミュニティコードのグループを対象とした「特別サポート」を合計 16 グループに対して実施している。「特別サポート」は、JCAHPC・NVIDIA 関係者による個別のサーヴェイ、主要カーネルの最適化などのサポートである。ただ、コード量が多かったり、アルゴリズムやデータ構造の変更が必要になる場合もあり、また将来の継続的なメンテナンスも考慮して、ミニキャンプ、講習会、相談会にも積極的に参加していただいている。

また、需要の多い、OpenFOAM 等のアプリケーションについては、NVIDIA とも協力し、OpenFOAM HPC Technical Committee¹⁴で進められている GPU 化の成果を活用する予定である。

さて、「第二軍の大輸送」である。

司馬遼太郎「坂の上の雲」は、筆者の少年時代からの愛読書で、40 年余りの間に繰り返し読んでいる。主人公の一人である秋山真之（日露戦争時の連合艦隊作戦参謀）の功績の一つに、「日本海海戦」と並んで「第二軍の大輸送」というのがあるらしいが、小説中には明確な記述がない。半藤一利「日露戦争史」によると、奥保鞏大将を総司令官とする「第二軍」の約 4 万人の将兵を乗せた 56 隻の大船団は、ロシア軍の旅順要塞のある遼東半島の先端近く、金州・南山付近に上陸した。ロシア旅順艦隊による攻撃、ロシア陸軍の水際攻撃を想定して、海軍の護衛艦の手配などもしていたのだが、結局抵抗らしい抵抗はなく、全員無事に上陸できたらしい。詳細はよくわからないのだが、とにかくこの大規模な輸送計画の立案者が秋山真之だった、ということらしい。

GPU へのアプリケーション移植のための活動を始めるにあたって、ふと「第二軍の大輸送」のことが頭をよぎった。そこで、今回の一連の活動を個人的に「第二軍の大輸送」と呼ぶことにした。「万全の準備をして、3,000 人の利用者全員が、一人の例外もなく、OFP-II の GPU を使えるようにする」ことがそのミッションである。「作戦名」はともかく、本件に関わる全ての人々がその目標を共有している。2022 年 10 月以降、これまで全ての活動が驚くほどスムーズに進んでいるのは、その賜物であると思う。

これから一年あまり、長い旅になるが、どうか利用者の皆様は、安心してともに進んでいただきたい。よろしくお願いします。

参考文献

- [1] A. Naruse, J.D. Trotter, J. Langguth, X. Cai, K. Nakajima, High resolution simulation of cardiac electrophysiology on realistic whole-heart geometries on Wisteria/BDEC-01 (Aquarius), 情報処理学会研究報告 (2022-HPC-187-15), 2022

¹⁴ [https://wiki.openfoam.com/High_Performance_Computing_\(HPC\)_Technical_Committee](https://wiki.openfoam.com/High_Performance_Computing_(HPC)_Technical_Committee)