

# Green500 連覇への道のり



安達知也 | (株) Preferred Networks

## Green500 の概要

Green500 は、バージニア工科大学の Feng 教授を中心とするグループによってメンテナンスされている、スーパーコンピュータのエネルギー効率のランキングである。関連するランキングとして TOP500 がある。TOP500 は LU 分解で連立一次方程式を解く LINPACK ベンチマークの実行速度を用いて 1 秒あたりの倍精度浮動小数点演算回数 (FLOPS) を競うランキングであり、近年では富岳が 4 期連続で 1 位を獲得している。Green500 は、TOP500 にランクインした上位 500 システムのスーパーコンピュータを対象に、FLOPS 値をベンチマーク実行中の平均電力で割った値、いわば電力あたりの演算性能を指標とするランキングである。筆者らは、2020 年 6 月、2021 年 6 月、2021 年 11 月の 3 回の Green500 で 1 位を獲得した。

スーパーコンピュータを評価する指標として絶対的な演算の速さだけでなく電力が重視されるのは、単純な電気代だけではなく冷却能力の観点でも限界があり、それによって演算速度が頭打ちになってしまうという危機意識に基づいている。1 つのスーパーコンピュータに供給可能な電力は数十 MW とされており (たとえば富岳は 30MW 弱の電力で TOP500 に登録されている)、分かりやすい基準として 1EFLOPS (1 秒あたり  $10^{18}$  回演算) を 20MW で実現する、す

なわち 50GFLOPS/W を達成するということが 1 つの大きな目標とされてきた。図-1 に 2007 年の初回 Green500 発表以降の 1 位の GFLOPS/W 値を示す。レギュレーションの変更やシステムの撤去等で若干のブレはあるがおおむね上昇傾向にあり、今年中には 50GFLOPS/W が達成されると予想ができる。

## MN-3 による Green500 連覇

MN-3 は Preferred Networks (以下 PFN) が構築した、MN-Core と呼ばれるアクセラレータを搭載したスーパーコンピュータである。近年のスーパーコンピュータは CPU とは別にアクセラレータを搭載して演算能力を高めることが多い。MN-3 においても演算能力の約 96% は MN-Core が担っている。MN-Core は PFN と神戸大学の牧野淳一郎

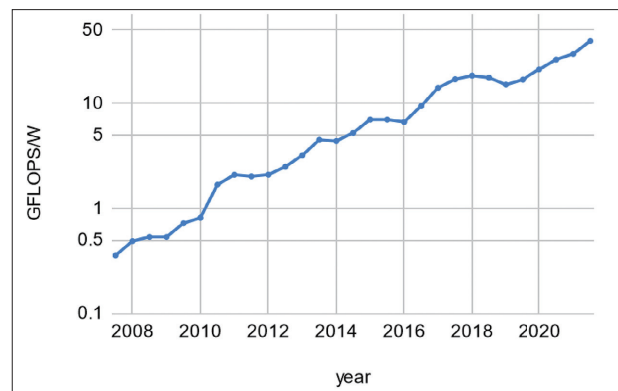


図-1 Green500 1位の推移

教授を中心とする研究グループで共同開発したプロセッサである。MN-CoreはPFNのビジネスの軸である深層学習を主なターゲットとしており、自社ビジネスを高速に実行することを主目的として開発された。また、今後広くエッジデバイスで深層学習などの大量の演算を要するAI処理を行うことを考えると省電力で高性能なプロセッサの実現は必須であり、その要素技術の実証という側面もあった。

高性能性・省電力性を実現するMN-Coreの大きな特長として、演算器以外の回路を極力減らしチップ面積に占める演算回路の割合を高めたという設計思想が挙げられる。これは牧野教授の研究グループで開発してきたGRAPEシリーズのプロセッサ、特にGRAPE-DRの思想を受け継ぐものである。汎用のプロセッサでは、多様なプログラムを効率良く実行するための制御回路（たとえばキャッシュ制御や命令間依存解析など）が無視できない面積を占めており、演算器に割けるチップ面積が小さくなってしまっているほか、演算に伴って制御回路も動作し消費電力が増えていた。一方でMN-Coreが扱うのは主に深層学習の演算であり、計算グラフという形で事前に演算内容が得られるため、ソフトウェア側で

あらかじめ解析することが比較的容易である。そこで、複雑な制御回路を搭載せずに極力ソフトウェア側で制御する設計とすることで、チップ面積に占める演算器の割合を高め、電力あたり演算性能も高めることができた。一般にプロセッサは半導体プロセスの微細化により消費電力が低下するが、12nmで製造されたMN-Coreが他社の7nmで製造されたプロセッサとGreen500で対等に渡り合っているのは、この設計思想によるところが大きいと考えている。

Green500への挑戦は2019年に始動した。LINPACKベンチマークの移植・最適化については後述するが、それ以外の重要な要素としてGreen500では当然電力計測が必要であり、高いサンプリング周波数で測定が可能な電力計の導入や、ベンチマーク実行と電力値を対応づけるソフトウェアの開発も並行して行われた。図-2にGreen500向けのベンチマーク実行および開発フローを示す。測定された電力はデータベースに蓄積され、時系列データ可視化ツールであるGrafanaでリアルタイムに消費電力の推移を観測することができるようになっている。また、ベンチマークプログラムからの測定開始・終了のイベント通知を契機として、1回のベンチマーク実行に対応する消費

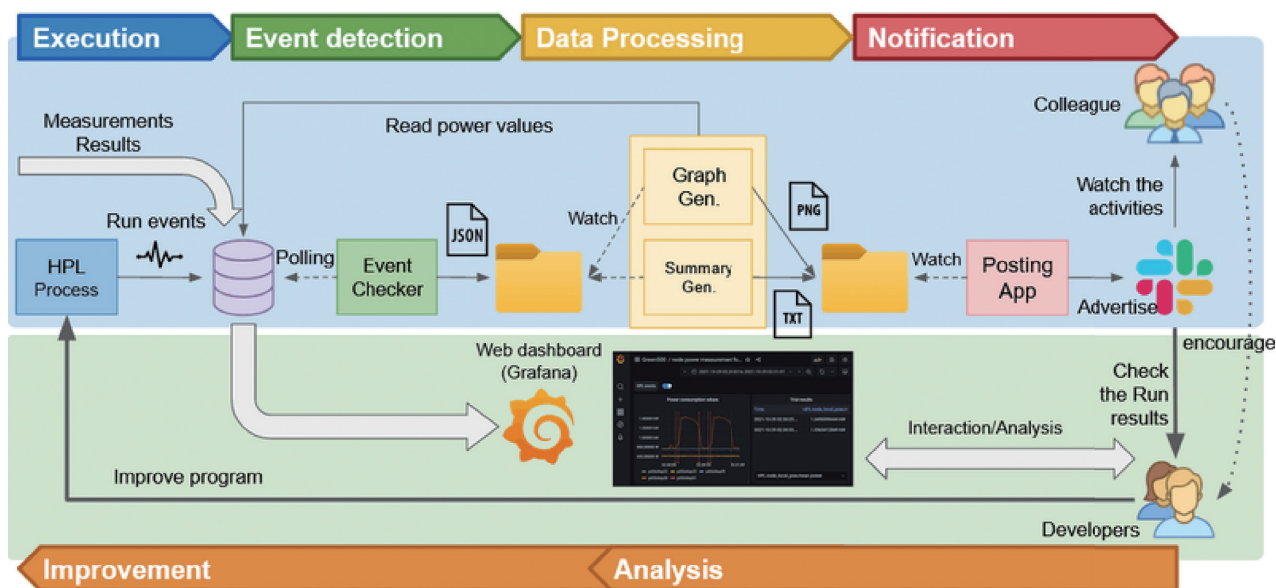


図-2 電力計測を組み込んだ開発フロー（図中アイコンは Icons8 (<https://icons8.com/>) より）

電力推移のグラフの生成や消費電力の積算が行われ、何 GFLOPS/W であったかが計算される。これらのグラフや値はベンチマーク実行直後に社内の誰でも閲覧可能な Slack に投稿され、開発チームはこの結果を参考にして次の最適化を進める。記録が更新された際などは同僚から称賛のコメントやリアクションが付き励みになる、といった具合である。

一方で肝心の LINPACK ベンチマークの移植は難航した。PFN としては初めてのプロセッサ開発でコンパイラなどの周辺ソフトウェアの整備が十分でなく、MN-Core 上で実行する命令をリアルタイムで生成する足回りのソフトウェアの開発から始める必要があった。結果として、2020年6月の Green500 時点ではベンチマークを動かすのがやっとという状態であり、最適化を十分に行うことができなかった。前述のとおり Green500 は TOP500 にランクインしたスーパーコンピュータを対象としたランキングなので、TOP500 にランクインするための最低限の演算性能が必要である。当初の予定では 32 ノードで計測した結果を提出する予定だったが、性能が足りないことが分かったため急遽 40 ノード実行に切り替えることとなった。そういったドタバタがありながらも、元々のプロセッサの電力あたり演算性能の高さに助けられて 2020年6月の Green500 では 21.108GFLOPS/W を達成し 1 位を取ることができた。

その後は本業の深層学習プログラムの開発の合間にプログラムの改良・測定が進められた。一度 1 位を取った時点で終わりとしなかったのは、改良の余地が十分にあると予測ができていたことに加え、プログラムの最適化ノウハウの蓄積、次世代以降のプロセッサ仕様へのフィードバックを見込んでのことである。この際、MN-3 を深層学習用の環境と Green500 用の環境との間で相互に切り替える手順が徐々に効率化され、週末などの短い期間であっても測定を行うことができるようになったのは開発の助けになった。

さて、電力あたり演算性能を上げるには、単純に消費電力を下げるか、プログラムを最適化して実行

時間を短くすることで演算器以外の回路が静的に消費している待機電力を下げるかが必要となる。電力を下げるためには、MN-Core のチップごとにばらつきのある動作可能電圧を見極めて個別に動作電圧を下げる等の施策を行った。プログラムの最適化に関しては、演算が動いている間にデータの移動や命令生成が行われるようにして演算が絶え間なく行われる状態を目指すということに尽きるが、MN-Core ではすべての命令のスケジューリングをソフトウェアであらかじめ行う必要があるため難易度が高く、段階的に最適化を進めていく形となった。

その結果、2020年11月の Green500 では 26.039 GFLOPS/W、2021年6月は 29.700GFLOPS/W と徐々に性能値を高めていった。2021年11月の挑戦では、ソフトウェアで生成する命令を工夫することにより MN-Core 内部の回路の電力消費を細粒度で制御する方法を新たに開発したこともあり、39.379GFLOPS/W と大きく性能値を上げることができた。初回挑戦時と比較すると電力あたり演算性能は約 2 倍に向上したことになる。

表-1 に 2020年6月から 2021年11月の Green500 ランキングのトップ 3 のスーパーコンピュータとその GFLOPS/W 値を示す。カッコ内はそのシステムが搭載しているアクセラレータの名称である。MN-Core が初登場で 1 位を獲得した 2020年6月の Green500 は NVIDIA 社の A100 と呼ばれるブ

表-1 最新 4 回の Green500 各 Top3

回	順位	スーパーコンピュータ (アクセラレータ)	GFLOPS/W
2020年6月	1	<b>MN-3 (MN-Core)</b>	<b>21.108</b>
	2	Selene (A100)	20.518
	3	NA-1 (PEZY-SC2)	18.433
2020年11月	1	NVIDIA DGX SuperPOD (A100)	26.195
	2	<b>MN-3 (MN-Core)</b>	<b>26.039</b>
	3	JUWELS Booster Module (A100)	25.008
2021年6月	1	<b>MN-3 (MN-Core)</b>	<b>29.700</b>
	2	HiPerGator AI (A100)	29.521
	3	Wilkes-3 (A100)	28.144
2021年11月	1	<b>MN-3 (MN-Core)</b>	<b>39.379</b>
	2	SSC-21 Scalable Module (A100)	33.983
	3	Tethys (A100)	31.538

ロセッサが初登場した回でもあり、以後 MN-Core と A100 がトップ争いを繰り広げることとなった。両者は行列演算器を大量に搭載したプロセッサであるという共通点はあるもののまったく異なるアーキテクチャである。それにもかかわらず、半年ごとの電力性能向上具合がほぼ同じで僅差で1位2位を取り合っていたというのは非常に興味深い。最新の2021年11月のランキングでは MN-Core が大きく突き放して連覇する結果となったが、これは前述の命令の工夫で消費電力を下げる手法の開発によるところが大きいと考えている。

2年間の記録向上の歴史は、MN-Core の使いこなし方を学習していくプロセスでもあった。現在、Green500

への挑戦を通じて得た知見をベースに、MN-Core の深層学習応用での利活用を進めている。さらに、次世代ハードウェアの開発も続けており、引き続き高効率な計算とこれを用いた応用の展開を行っていきたい。

(2022年2月25日受付)  
(2022年4月5日 note 公開)

安達知也 (正会員) [adachi@preferred.jp](mailto:adachi@preferred.jp)

2010年東京大学大学院情報理工学系研究科コンピュータ科学専攻修士課程修了。富士通(株)でのスーパーコンピュータ向けソフトウェア開発を経て、2017年より現職にてプロセッサ開発、アプリケーション開発に従事。

