

## スパコンの節電運用について

岩橋建輔<sup>#</sup>

自然科学研究機構 分子科学研究所

### 概要

2022年度まで電力単価が上昇しており、2023年度当初の単価が維持された場合、計算科学研究センターの電気代予算を超過することが確実であった。スパコンの消費電力を減らす必要があったが、具体的な方策には様々なものが考えられた。運用ノード数は減らさない方針が示されたため、CPUのクロックをどれくらい落とすかがポイントとなった。その判断材料として、CPUの動作モードを変えて分子動力学計算の計算時間とノードの消費電力を調べた。

その後、計算に使われているCPUのコア数-GPU数と消費電力の関係の可視化と月間消費電力量の予測を行なえるようにしたが、様々な要因により消費電力に揺らぎが生じていることも紹介する。

### 1. 過去4年の電力の状況

#### 1.1 電力単価

最近4年間の電力単価の推移をグラフにしたのが図1である。2021年から徐々に燃料価格が高騰しウクライナ侵攻の影響によりさらに高騰したことが電力単価高騰の一因である。2023年4月をピークに下落に転じている。2023年5月から再生可能エネルギー発電促進賦課金単価が3.45円/kWhから1.40円/kWhに引き下げられ、電力会社による軽減単価が2023年6月から10月までの間2.09円/kWhが適用され、石炭価格等が下落した影響である。

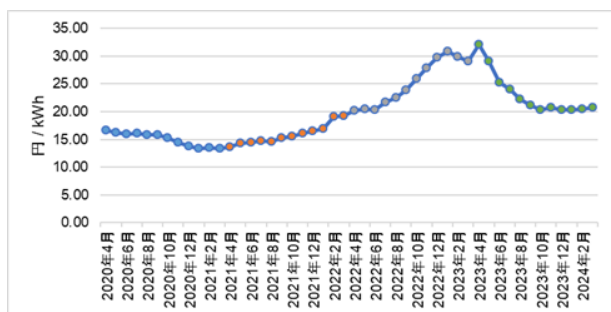


図1. 最近4年間の電力単価の推移

#### 1.2 計算科学研究センターの使用電力量と電気代

図2は最近4年間の使用電力量の推移である。2022年9月までは旧スパコンが稼働しており夏場には500MWhを使用することもあった。2022年10月から2023年1月まではシステム更新のため、ファイルサーバーや仮想サーバーのみが稼働しており、使用電力量は大きく落ち込んでいる。2023年2月から運用開始した現スパコンは消費電力を抑えるため水冷システムを採用している。それにより全ての演算サーバーが使用されていた2023年3月でも444MWhであった。

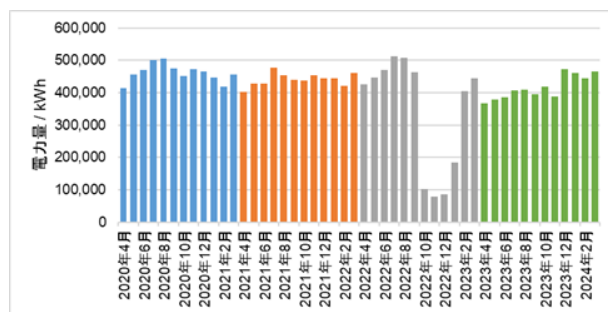


図2. 最近4年間の使用電力量の推移

しかし、電力単価が最高値付近であったため月額の電気代が約1,300万円となった。そのため2023年度の運用では、電力単価が維持もしくはさらに上昇した場合も想定しなければならなかった。図3は最近4年間の電気代の推移である。2023年度だけ積み上げグラフにしているが薄い色で積み上げられた方は4月の単価が維持された場合のものであり、年額にすると約1億6,000万円となっていた。幸い電力単価が下がったため、実際の電気代は濃い色の方であり、年額では1億1,500万円程度となっているが、2023年4月の段階では電気代の予算超過が予想されたことから、使用電力量削減を進める必要があった。

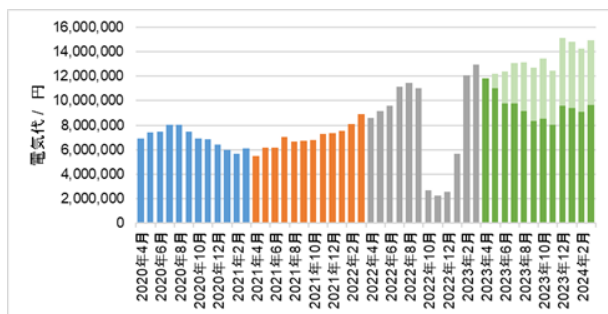


図3. 最近4年間の電気代の推移

### 2. 使用電力量削減

```
# ipmitool dcmi power reading | grep Instantaneous
Instantaneous power reading:      246 Watts

# grep MHz /proc/cpuinfo | head -3
cpu MHz      : 2445.300
cpu MHz      : 2445.300
cpu MHz      : 2445.300
```

図4. 電力値取得とCPUクロック取得の例

表1. CPUの動作クロックとアプリケーション動作時の消費電力および計算時間

	アプリケーション	CPU (MHz)	消費電力 (Watts)	計算時間 (sec)
Turbo boost on OnDemand 最大 2.68 GHz	-	1511.29	250	-
	gromacs	2679.36	687	627.351
	HPL	2453.23	724	-
Turbo boost off OnDemand 最大 2.45 GHz	-	1515.29	251	-
	gromacs	2445.17	639	657.377
	HPL	2445.26	710	-
Turbo boost off 2.45 GHz 固定	-	2445.18	255	-
	gromacs	2445.19	634	652.953
	HPL	2444.37	709	-
Turbo boost off 2.0 GHz 固定	-	1999.06	255	-
	gromacs	1996.12	513	767.336
	HPL	1999.00	629	-
Turbo boost off 1.5 GHz 固定	-	1499.98	249	-
	gromacs	1499.39	409	996.428
	HPL	1499.25	462	-

表2. Gromacs 実行時の消費電力の削減率と計算時間増加率

CPU (MHz)	消費電力 (Watts)	経過時間 (sec)	消費電力の削減率	計算時間増加率
Turbo boost on 2.68 GHz (OnDemand)	687	627.351	100%	100%
Turbo boost off 2.45 GHz (OnDemand)	639	657.377	93%	105%
Turbo boost off 2.45 GHz (固定)	634	652.953	92%	104%
Turbo boost off 2.0 GHz (固定)	513	767.336	75%	122%
Turbo boost off 1.5 GHz (固定)	409	996.428	60%	159%

表3. CPUクロックの制御方法毎の消費電力と計算時間

	アプリケーション	CPU(MHz)	消費電力(Watts)	計算時間(sec)
BIOS 制御 OnDemand 最大 2.45 GHz	-	2445.30	246	-
	gromacs	2445.30	518	44.611
	gaussian	2445.30	509	1347
acpi_cpufreq 制御 OnDemand 最大 2.45 GHz	-	1500.00	248	-
	gromacs	1500.00	544	44.653
	gaussian	1500.00	539	1355

## 2.1 削減の方針について

現スパコンはエアコンを使わない水冷システムであり、以前のような空調消費電力削減対策[1]を行えない。そのため、演算サーバーの稼働台数を減らすか、CPUの動作クロックを下げた運用するかの選択肢しかなかった。計算科学研究センターで議論した結果、稼働台数を減らすことの方がユーザーへの影響が大きいとの結論になりCPUの動作クロックを下げた運用する方針となった。

## 2.2 CPUクロックをどれくらい下げるか

一般にCPUクロックを下げると消費電力は低くなるが計算速度が増大する。定量的に測定する必要あり、導入したスパコンの演算サーバーはDCMI (Data Center Manageability Interface) 仕様に準拠しているのでIPMI (Intelligent Platform Management Interface) を通じてその瞬間の使用電力を計測した。またCPUクロックに関してもLinuxの/proc/cpuinfoから得ることができる。具体的な取得方法は図4にあるようにそれぞれipmitoolコマンドとgrepコマンドで得られる。

準備が整ったので比較的計算効率のよいプログラムを使って消費電力と計算時間を測定した。使用したプログラムは分子動力学シミュレーションの計算で使われるGromacsとスパコンの演算性能を計る際に使われるLINPACKの実装のHPLである。HPLについてはCPUに負荷をかけることを目的としたので計算時間については測定していない。また、アプリケーションを動かしていない状態での消費電力も計測した。CPUクロックはターボブーストが有効でOnDemand (計算負荷が高いときに自動的にCPUクロックを高くするモード)のもの、ターボブーストを無効にしてCPUクロックを2.45 GHz、2.0 GHz、1.5 GHzに固定したものが設定でき、合計4種類のCPUクロックで計測し、それらを表1にまとめた。表1の結果から消費電力の削減率と計算時間の増大率を求めたものが表2である。

2023年4月の段階ではCPUクロックの設定変更によりシステム全体でどれくらいの電力が削減できるか不明であったため、表2の結果からとりあえずターボブーストをオフにして動作クロックは最大の2.45 GHzで運用し、CPUの消費電力は7%減、計算時間は5%増の目論見で様子を見ることとした。

この頃別件でLoad Averageが実装されているコア数の128を大きく超えていることが判明し、原因を追究したところCPUクロックを制御するLinuxのカーネルモジュールのacpi\_cpufreqが原因であった。そのため、acpi\_cpufreqを使わない方法としてCPUクロックの制御をBIOSに任せる方法があることが分かり、同様に消費電力等を計測した。この計測時にはGromacsだけでなく量子化学計算に使われるGaussianも用いた。その結果を表3にまとめた。BIOS制御のOnDemandのときアイドル時のCPUクロックは2.45 GHzであるが実際の消費電力は1.5 GHzのときと変わらなかった。また、計算時間についてもBIOS制御の方が若干短かったため、2023年6月からBIOS制御のターボブースト無効の状態に運用した。

2023年後半になると電力単価がだいぶ下がってきたので電気代予算内に収まる目途が付き始めた。そこで2023年12月からBIOS制御でターボブースト有効な状態にして運用している。

## 3. CPUやGPUの使用数と電力の関係把握

現スパコンの消費電力はマルチメーターの値をネットワーク経由で入手可能となっている。一方、冷却塔や既設の空調機器の消費電力については分電盤にIEEE1888の機器を取り付け、こちらもネットワーク経由で入手できるようになっている[1]。

このように入手した電力値はGrafanaを用いて可視化しているが、CPUやGPUの使用状況と関連付けられていないため、月間の電力量を予測することは不可能であった。そこでCPUやGPUの使用状況と消費電力の相関を得るために解析するプログラムを作成して可視化(図5)した。

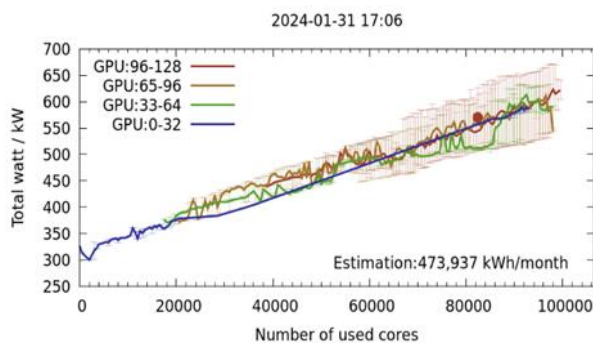


図5. 使用CPU、GPUと消費電力の関係

図5から明らかになったことは演算サーバーで計算していない状態でも300kW使用している。これはファイルサーバー、InfiniBand、仮想サーバー等の電力である。当初演算サーバーのターボブーストをオフにすることでCPUだけで7%の削減としていたのは実はシステム全体では約3%の削減を意味していた。しかし、実際には10%以上電力を削減しており、Gromacsに比べて効率の悪いアプリケーションではターボブーストによる加速が得られなくなった影響による電力削減が大きかったと推測される。

また、月間電力量予測値は収集した値の積算値に経験上1.1倍したものである。これらは居室で使用した電力を考慮していない影響によるものと思われるが、実際に確かめるのは困難である。

## 4. まとめ

現スパコン導入2か月後に全体の消費電力を抑制する運用を模索しなければならなかった。なるべく論理的にCPUクロック等の調整を行ったものの、実際には想定漏れの事象が多数あり、偶然にも適度な節電となったのは幸運であった。

## 参考文献

[1] 松尾純一、他、”計算科学研究センターにおける空調消

令和 5 年度高エネルギー加速器研究機構技術研究会、2024 年 3 月 7～8 日  
<https://www2.kek.jp/engineer/tsukuba/giken/2023/>

費電力削減対策結果”平成 27 年度高エネルギー加速器研究機構技術研究会報告集 p164-172.