

数値計算環境の構築について

To Construct a Computing Environment for Numerical Simulation

嶋崎 真一*

Shin-ichi SHIMASAKI

概要

数値計算の環境を整えようとしたときに、一般向けの PC と、科学技術計算用の HPC (High Performance Computer) の 2 つの選択肢がある。最近では PC の低価格化が著しく、Intel の高性能 CPU を搭載した計算機が 10 万円台で買えてしまう。一方で、科学技術計算用と銘打った計算機を探すと、ちょっとしたスペックでも数十万円、少し良い構成にするとすぐに百万円単位の予算が必要になる。この差が何に由来しているのかを調査し、計算機を選択する際の指針について簡単にまとめた。

Keywords: HPC, 数値計算, ソフトエラー, ECC メモリ

1. 中性子線とソフトエラー

香川高専高松キャンパスの学生用玄関には大型の霧箱が設置されており、霧の白い軌跡が常時現れては消えることが観察できる。この軌跡は放射線が通過した際に生じるものであり、空間を常に放射線が飛び交っていることが実感できる。放射線には幾つかの種類があり、放射線ごとに軌跡の様子が異なっていて、じっと見ているだけで楽しい。

放射線の中には中性子線と呼ばれるものがあり、海拔 0 m では 1 cm^2 , 1 hour あたり 13 個程度飛来している¹⁾。これだけの頻度があれば、霧箱で容易に観察されても良さそうなものだが、中性子は名前の通り電氣的に中性で、物質を透過する能力が非常に高い。そのため、中性子線は霧箱と相互作用を起こすことなく透過してしまう。残念ながら霧箱ではその軌跡を観察することはできない。

この中性子線が運悪く計算機のメモリに衝突し、ごく稀にノイズが生じてメモリの内容が書き換えられてしまうことがある。これがソフトエラーと呼ばれるも

のでありⁱ、再起動すれば何事もなかったように復帰する。ソフトエラーは、どのタイミングでどの場所に起こるか予測できず、再現性がない。

ソフトエラーの頻度については、緯度や高度・太陽活動の影響などが知られている¹⁾。高度や緯度が高いところにある場合は、宇宙線が多いためにソフトエラーの頻度が上昇する。富士山山頂では海拔ゼロの地点よりもソフトエラーの確率が 17 倍高くなるし、高高度を飛ぶ旅客機内では確率は 300 倍も高くなる。また、太陽活動が低下すると太陽風によって遮られていた宇宙線が増えるので、ソフトエラーが全世界的に増加することになる。

使用しているメモリの配線幅が微細化されると、その配線が細くなった分だけソフトエラーに対する耐性が少なくなる一方で、宇宙線がヒットする回路面積が減少する。この二つの効果が相殺しあって、結局ソフトエラーの頻度は世代によってそれほど変化していない。一方で、システムの経年劣化によってソフトエラーが増えると言われており、使用開始後 10~18 ヶ月

ⁱ ソフトウェアのエラーという意味ではないので注意。一方、再起動しても復帰できないような、物理的な回路の故障はハードエラーと呼ばれる。

* 香川高等専門学校 機械電子工学科

でエラーの発生率の増加が始まるとの報告がある²⁾。

Daniel らは実際に数値計算を行っている CPU に強力な中性子線を照射し、CPU のキャッシュメモリにおけるソフトエラーの加速試験（自然界に換算して 57,000 年に相当）を行い、その結果を報告している³⁾。実行している科学技術計算の種類によってエラーの発生率が異なったり、アルゴリズムを変更することによってエラー率を低減させることが示されており、大変興味深い。また、彼らの実験結果からすると、大規模な計算機（日本で言えば「京」などに相当）では、1 日あたり 1 回程度は CPU のキャッシュメモリに深刻なエラー（検出はできたが訂正ができないタイプのエラー）が生じていることになる。昔、某大学がメモリ訂正機能のない市販の PC 約 1,000 台を組み合わせるスーパーコンピュータを作ろうとしたが、ブート途中のわずかな間にソフトエラーで倒れるという状態で使いものにならなかったという⁴⁾。

以上の情報を総合すると、我々が通常使用している計算機においては、大雑把にいつに 1 回程度ⁱⁱ⁾はソフトエラーが生じていると考えて良さそうである。計算機は意外にドジっ子なのだ。

ソフトエラーが生じたとしても、ビット反転した箇所がたまたま未使用領域であれば、計算機への影響はない。また計算機がハングアップしてくれば、計算は中断こそしてしまうが、それ以上の問題にはならない。厄介なのは、データの一部が化けてしまったことに気がつかないまま作業を進めてしまうことである。化けたデータに基づいて誤った結論を出してしまうと取り返しがつかないし、何か変だと思ってもソフトエラーには再現性がないので、どこが悪かったのか検証すらできない。こうなるとドジっ子どころの話ではなくなる。Geist 氏はこう問いかけている⁴⁾：

You wouldn't want to fly in an airliner designed using such a calculation. Nor would you want to certify a new nuclear reactor based on one.

このような計算を使って設計された旅客機には乗りたくないだろう。もしくは、このような計算に基づいて新しい原子炉を保証したくないだろう。

ⁱⁱ⁾ 既に説明したように、ソフトエラーの頻度は、緯度・高度・太陽活動の強度・マシンの経年劣化の度合いやメモリ搭載量などに依存している。この頻度は 0.1 倍～10 倍程度の幅をもって考えて欲しい。

2. ソフトエラー対策（ECC メモリ）

現在のところ、最も普及しているソフトエラー対策は、ECC メモリを使用することである。ECC メモリとはエラー訂正機能付きのメモリ Error Check and Correct memory（もしくは Error Correcting Code）のことである。メモリにエラー訂正用の冗長データを付加して、これを用いてソフトエラーによってビット反転したデータを自動で検知し、訂正することができる。ECC の仕組みや動作については、ネットを少し調べただけでもたくさんの情報が出てくるので、本稿では省略する。

ECC が必要とされるケースは以下のとおり：

- 財務計算や一部の科学技術計算のように、数値の信頼性が絶対的に重要な場合。
- 長時間連続運用する場合。数値計算を長時間連続で回しているような使い方をしていないと、その間にソフトエラーに遭遇する可能性が高くなる。
- 大量のメモリを積んでいる場合。メモリをたくさん積んでいると、それだけソフトエラーに遭遇する可能性は高くなる。

逆に考えると、以下のようなケースでは無理に ECC を使用しなくても良い。

- たまにハングアップしても、再起動すればそれで済ませられる場合。
- マシンの大半の時間はアイドルで占められている場合。
- メモリがそれほど多く無い場合。

通常の事務的な用途では ECC は不要だろう。出力画像に多少のノイズがあっても許容できるのであれば、長時間にわたるレンダリングやエンコードを行う計算機にも ECC は必要ない。科学技術計算でも、多少ビット反転が生じても全体的な結果には影響がないようなケースⁱⁱⁱ⁾であれば、問題はない。

なお通常の ECC によって訂正できるのは 1 bit のエラーのみである。1 bit のエラーは ECC が自動的に訂正し、ログを残してそのまま計算機は走り続ける。

ⁱⁱⁱ⁾ モンテカルロ法によるシミュレーションや、ディープ・ラーニングなどが該当するだろう。また近年においては、多少のエラーが発生することを前提にして、許容誤差を設定して計算機を運用することも考えられている³⁾。

減多にないことであるが、同一メモリ上で2 bit が同時に反転してしまうことがある。この場合、エラーを訂正することはできないが、エラーの検出は可能である。最近の Linux では訂正不能なエラーが検出された場合、即座にカーネルパニックを発生させることができるようになってきている。3 bit 以上の場合は、検出すらできない可能性が出てくる。こんな時は、運が悪かったとして諦めるしかない。

近年では複数のメモリセルの値が同時反転してしまう現象や、並列化された計算機の複数の論理ノードに同時にエラーが生じてしまう現象が顕在化してきており⁵⁾、それらに対する対策が盛んに議論されている。ECC を拡張してマルチ bit のエラー訂正ができるもの(商品名: Advanced ECC, ChipKill など)はハイエンドのサーバに多く採用されているし、さらに心配性の人はメモリを二重にして、ECC で訂正不能なエラーが出た時はもう一方の内容を取り出して使用したりする(ミラーリング)。当然、メモリは搭載量の半分しか使えない。二重で足りなければ、三重にする。他にもメモリ内容の保護のために様々な技術が存在しているが、それらは総じて恐ろしく高価である。間違いが絶対に許されない銀行の勘定系や、宇宙線の影響が深刻な航空機や宇宙探査機で採用されているようだ⁶⁾。

なお、先にも述べたようにソフトエラーを引き起こす宇宙線は、一般に中性子線(特に高速中性子)である。中性子線は遮蔽するのが非常に難しく、水やポリエチレン、カドミウム、コンクリートなどの層を複合的に組み合わせて遮蔽される⁷⁾。遮蔽によってソフトエラーを低減させるのは現実的ではない。

3. ECC が使用できる計算機

本稿執筆の2017年度の時点では、Intel社のCPUでECCを使用できるのはXeon系列のみとなっている。Core i5, i7, i9ではECCは使用できない。かつて2006年にWindows Vistaがリリースされた時に、MicrosoftがECCメモリを必須にしようと画策したが、コストアップにつながるとしてメーカーが反対して、上手くいかなかったことがあった。それ以来ずっと、一般向けのPCはnon-ECC、サーバ向けはECCと棲み分けられている。

Xeon そのものが高価であることに加えて^{iv}、Xeonを選択した時点で、使用できるマザーボードやメモリなどの選択肢が少なくなり、価格が跳ね上がる。ある程度のメモリ容量を搭載した計算機を購入しようとするなら、簡単に100万円を超えてしまうだろう。

頑張つて安くあげようとするなら、Core i3やXeon E3でECCが使えるマザーボードで組み立てることも可能らしい⁸⁾。しかし、自作で安定した計算機を組み上げるのはかなりのノウハウが必要であるし、そもそも公式にはサポートされていない。信頼性の向上を求めてECCを使用するのに、信頼性を確保するのが難しい自作機を選択するというのは無理がある。よほど腕に自信があるのではない限り、業務に使用するのであれば、信頼のできるベンダーから既製品を購入すべきである。

一方、最近発売されたAMD社のCPUであるRyzenでは、コンシューマ向け製品でECCの動作を検証していないが、CPUでの機能自体は有効であり、マザーボードとBIOS次第で対応可能とされている⁹⁾。これが本当なら、大容量のECCメモリを積んだマシンが安く調達できる可能性がある。今後の実績を待ちたい。

また次世代のメモリ規格であるDDR5では、メモリそのものにECCが実装されることになっている¹⁰⁾。この製品が普及すれば、一般用途のPCでもECCが使用できるようになるだろう。この製品が実際に登場するのは2020年頃と言われている。

4. では、どちらを購入すべきか?

Dellあたりで適当に検討すれば、高速なCoreシリーズのCPUを搭載したnon-ECCのマシンが安価に入手できる。並列数が少ない計算であれば、Xeon搭載の計算専用機より速いことも珍しくない。ソフトエラーが起りうることを前提にして、相応の手間をかけることを厭わないのであれば^v、あえてnon-ECCの計算機で済ませるという判断があっても良い。一方で、ソフトエラーが受け入れられないのであれば、現状ではXeon + ECCの計算機にするしかない。

^{iv} 最も高価なものだと、CPU単体で100万円を優に超える。基本的に業務用であり、CPUに投資してもそのコストを業務で回収できることを前提として購入するものである。

^v 例えば、重要な計算は二回行う、長時間の計算は行わないと割り切って運用する、など。

月並みな結論だが、予算／計算の重要性／費やすことのできる時間や手間などの費用対効果を勘案して、選択することになるだろう。

5. 当研究室の計算環境

筆者の研究室では、数値計算用に以下の計算機を用意している。主な用途は流体計算と大容量のメモリを要する画像処理である。

計算機名: weber

マザーボード: Supermicro 社製 X10DAI

C P U: Xeon E5-2667v3 (20 MB Cache, 3.20 GHz, 8 Cores) × 2 CPU (合計 16 コア)

メモリ: 128 GB (16 GB DDR4-2133 ECC Registered × 8)

OS: Ubuntu 16.04 LTS

ベンダー: HPC システムズ株式会社

導入時期: 2015 年 8 月

CPU の E5-2667v3 は、導入当時の Xeon の中で比較的動作周波数の高いものである。この計算機では主に OpenFOAM を用いて流体計算を行っているが、それほど並列性が良くない^{vi}。多少コア数を減らしてでもシングルの性能を上げたかったので、この CPU を選択した。メモリは当然 ECC で、製品出荷前にベンダーによって 58 時間の負荷試験が行なわれている。この構成で、購入当時の価格は約 150 万円であった。業務に使用するための計算機としては十分にリーズナブルであろう。他にも無停電電源装置や 10 GB Ethernetなどを付けている。

他にも non-ECC の計算機が複数台あるが、これらは全てその場で終了するような小規模な計算にしか使わない。手元の non-ECC の計算機でテストを繰り返した後、本番の計算を weber で行うという使い方をしていく。さらに大規模な計算が必要な場合は、他機関の大型計算機を使うようにしている。

これまで筆者は、どちらかというと実験がメイン、計算をサブで研究を行ってきたが、高専に赴任してきたのをきっかけに、計算のウェイトを増やしつつある。

^{vi} 小規模な計算 (約 40 万メッシュの流体計算) の場合、並列数 8 と 16 のときの並列化効率の実測値はそれぞれ 0.8 と 0.6 程度であった。なお OpenFOAM の並列化手法はフラット MPI である。

それにともない、少しずつ計算環境を整えてきた。これから計算を始めようという人に、本稿が少しでも参考になれば幸いである。

参考文献

- [1] JEDEC Solid State Technology Association, “Measurement and Reporting of Alpha Particle and Terrestrial Cosmic Ray Induced Soft Errors in Semiconductor Devices”, JEDEC Test Standard No. 89. 2001-08.
- [2] Schroeder, Bianca; Pinheiro, Eduardo; Weber, Wolf-Dietrich. “DRAM Errors in the Wild: A Large-Scale Field Study”. Proceeding of SIGMETRICS/Performance’09. Seattle, 2009-06-15/19, Association for Computing Machinery (ACM). 2009, p. 193–204.
- [3] Oliveira, Daniel; Pilla, Laércio; DeBardeleben, Nathan; Blanchard, Sean; Quinn, Heather; Koren, Israel; Navaux, Philippe; Rech, Paolo. “Experimental and Analytical Study of Xeon Phi Reliability”. SC ‘17 Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis. Denver, Colorado, 2017-11-12/17, Association for Computing Machinery (ACM). 2017, Article No. 28.
- [4] Geist, Al. “How To Kill A Supercomputer: Dirty Power, Cosmic Rays, and Bad Solder”. IEEE Spectrum. 2016-02-23. <https://spectrum.ieee.org/computing/hardware/how-to-kill-a-supercomputer-dirty-power-cosmic-rays-and-bad-solder>, (参照: 2017-08-31).
- [5] 伊部英史; 鳥羽忠信; 新保健一; 上菌巧; 谷口斉. “環境放射線による電子装置のソフトエラー・障害対策の現状と取り組み”. 日立評論. 2014, vol. 96, no. 07-08, p. 496–501.
- [6] Schmidt, Andrew G.; French, Matthew; Flatley, Thomas. “Radiation Hardening by Software Techniques on FPGAs: Flight Experiment Evaluation and Results”. Aerospace Conference 2017 IEEE, 2017, p. 1–8.
- [7] 東北大学環境・安全委員会; 原子科学完全専門委員会. 放射線とアイソトープの安全取り扱い—東

北大学放射線新規取扱者講習会テキスト 講義編.
2011, p. 102.

- [8] 山県. “MSI 「E3 KRAIT GAMING V5」 —ECC
メモリーが使える白黒仕様のゲーミング
マザーが MSI から発売”. ASCII.jp. 2016-
07-13. [http://ascii.jp/ele/000/001/193/
1193686/](http://ascii.jp/ele/000/001/193/1193686/), (参照: 2017-08-31).
- [9] Campbell, Mark. “AMD Confirms That Ryzen
Supports ECC Memory”, Overclock3D. 2017-
03-02. [https://www.overclock3d.net/news/
cpu_mainboard/amd_confirms_that_ryzen_
supports_ecc_memory/1](https://www.overclock3d.net/news/cpu_mainboard/amd_confirms_that_ryzen_supports_ecc_memory/1), (参照: 2017-08-31).
- [10] 福田昭. “次世代のサーバー／ハイエンド PC 向け
DRAM 「DDR5 メモリ」”. PC Watch. 2017-08-
23. [http://pc.watch.impress.co.jp/docs/
column/semicon/1076835.html](http://pc.watch.impress.co.jp/docs/column/semicon/1076835.html), (参照: 2017-
08-31).