

スーパーコンピュータ「富岳」における故障ノードに関する研究

発表者: 坂上創哉
所属: 三輪研究室
指導教員: 三輪 忍 准教授

1 はじめに

スーパーコンピュータの運用では、大量の計算ノードを扱うため、月に数十件の障害が発生することが報告されている [?]。計算ノードはマザーボードに固定されており、一つの計算ノードが故障した場合でも、CMU 単位での交換が必要となる [?]。このような障害は運用コストの増加やユーザーの利便性の低下を招くため、可能な限り減少させる必要がある。本研究では、「富岳」における故障ノードのログを分析し、故障の傾向や相関関係を明らかにすることを目的とする。

2 研究背景

2.1 スーパーコンピュータ「富岳」

「富岳」は、理化学研究所と富士通が共同開発した日本を代表するスーパーコンピュータである。2020 年 6 月に稼働を開始し、2020 年 11 月には TOP500 ランキングで世界 1 位を獲得した [?]。「富岳」は、ARM アーキテクチャを採用した A64FX プロセッサと、最大で 158,976 ノードの計算ノードを持つ [?]。富岳は、一般的な本用ソフトで動き、シミュレーションやビッグデータ、AI など幅広いアプリケーションで高性能を発揮するように設計されている。

2.2 ノードの故障

富岳の試行運用開始である 2020 年 4 月 1 日から 2024 年 10 月 31 日までの期間に記録された CMU の障害ログは、全 1,873 件でありこのうち、CPU まで特定されている障害件数は 1,579 件であると報告されている。[?]。また、富岳以前に理化学研究所で運用されていたスーパーコンピュータ「京」では、2012 年 9 月 28 日から 2015 年 6 月 30 日までの稼働をしていた。月別の CPU の故障率は 0.004% 程度の故障率で概ね安定している。[?]。0.004% とは毎月 3.3 件の故障であり、9 日に 1 個程度である。

2.3 電力ばらつきと故障

一般的なスーパーコンピュータは、同一使用の計算ノードがネットワークを介して接続されているため、これらの計算ノードは、与えられたアプリケーションに対して同一の消費電力を示すと期待されている。しかし、実際には半導体チップの製造上のばらつきが存在するため、同一のアプリケーションを実行してもそれぞれ異なる消費電力を示すことが報告されている。文献 [?] では、CPU の故障要因の一つとして、温度を挙げている。CPU の温度は、消費電力に依存しており、消費電力が高いほど CPU の温度も高くなるため、消費電力のばらつきが CPU の故障に影響を与えられられる。消費電力と故障率の相関関係は、文献 [?] で調査されており、消費電力のばらつきと故障率の関係性は見られなかった。

3 先行研究

文献 [?] では、ロスアラモス研究所が 10 年間に渡り収集したデータを用いて、HPC システムにおける故障の予測因子やその関連性を詳しく分析している。HPC システムにおける故障の傾向を広範に調査し、時間的・空間的な相関、故障原因の性質、およびノードの使用状況や環境要因が、故障の発生に大きな影響を与えることを明らかにしている。また、文献 [?] では、富岳における電力ばらつきがノード故障率に与える影響を調査している。富岳の電力ばらつきと故障の関係について、故障ノードの平均消費電力とノード全体の平均消費電力の差はわずか 0.3W であり、消費電力の差はほとんど見られず、電力ばらつきと故障率の関係性はあまり見られなかった。

4 研究目的

本研究では、電力ばらつきと故障の関係性が見られなかったことから、ノードの故障率には、利用率や実行されるジョブの種類も影響を与えられられる。富岳の故障ノードのログを分析し、故障の傾向や相関関係を明らかにすることを目的とする。

5 今後の予定

- 富岳の故障ノードのログを分析し、故障の傾向や相関関係を明らかにする。
- 故障ノードのログをもとに、故障の予測モデルを構築する。
- 故障の予測モデルを用いて、富岳の運用における故障率を低下させるための提案を行う。