

## スーパーコンピュータ「富岳」におけるノード故障の分析

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は多数の計算ノードによって構成されていることから、その運用中は多数 (数十件/月程度) の障害が発生することが報告されている [1]。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、我が国のフラッグシップスーパーコンピュータである「富岳」における故障ログを分析し、ノード故障の傾向や要因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 [3]。「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 [5]。富岳は 432 個のラックで構成され、384 ノードと 192 ノードの 2 種類が存在する。各計算機ラックにはシェルフと呼ばれる塊が格納されており、384 ノードのラックには 8 つ、192 ノードのラックには 4 つのシェルフが格納されている。1 つのシェルフには BoB (Bunch of Blades) が 3 つ取り付けられ、それぞれの BoB には 8 枚の CMU (CPU Memory Unit) が搭載されている。更に、1 枚の CMU には CPU が 2 つ取り付けられている。また、計算タスクのみを実行するノードは、アクティブ光ケーブル (AOC) ノードと non-AOC ノードの 2 種類に分けられる。AOC ノードはアクティブ光ケーブルに電力を供給し、non-AOC ノードは電力を供給しない。富岳は、一般的な汎用ソフトで動き、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 [4]。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。図 1 に示すように、月別の CPU の故障率は高負荷時を除いて 0.004% 程度で概ね安定している [2]。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、同一仕様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ことが期待されている。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

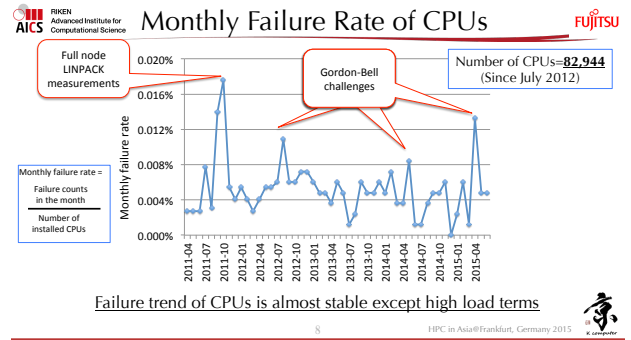


図 1: 「京」の月別 CPU 故障率 (文献 [2]) より

表 1: non-AOC ノードにおける消費電力 (文献 [4]) より

種類	数量	平均消費電力 (W)
全 non-AOC ノード	12,672	142.1
全故障ノード	107	141.8
要因: CPU	37	141.9
要因: メモリ	70	141.8

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 [4] 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 の温度も上昇するため、消費電力のばらつきが CPU の故障に影響を与えると考えられる。文献 [4] では、消費電力と故障率の相関関係を調査した結果、消費電力のばらつきと故障率の間に明確な関係性はみられなかったと報告されている。図 2 は富岳における消費電力と故障ノードの分布を示しており、ひし形で強調された点が故障したノードを表している。故障ノードは、消費電力に関わらず、ほぼ一様に分布していることが確認できる。また、表 1 は、富岳の non-AOC ノードにおける消費電力を示している。全体の平均消費電力は 142.1W であり、故障ノードの平均消費電力は 141.8W である。故障ノードの平均消費電力は全体の平均消費電力とほぼ同じであり、故障ノードの消費電力は全体の平均消費電力とほぼ同じであることがわかる。

## 3 先行研究

文献 [1] では、ロスアラモス研究所が 10 年間に渡り収集した大規模な運用データを用いて、HPC システムにおける故障の予測因子やその関連性を詳細に分析している。故障の発生傾向を時間的・空間的に広範囲に調査し、ノード故障の発生要因と他のノードへの影響範囲について包括的な知見を提供している。特に、HPC システム内部の物理的な構成部品が故障し、ノード単位でのが障害を引き起こすことががあると報告されている。故障の原因としては、電源品質や冷却系の問題などの環境要因が大きく関与していることが示唆されている。さらに、同一ラッ

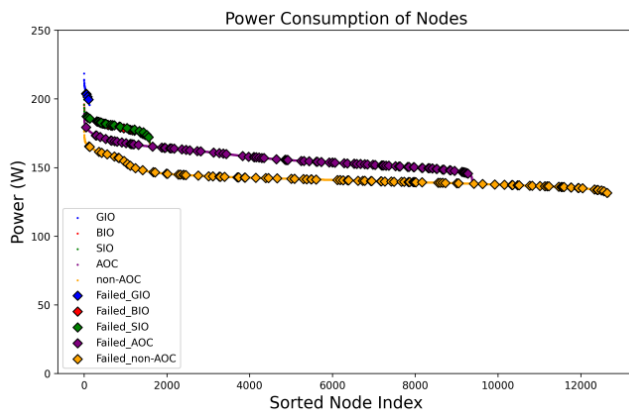


図 2: 富岳における消費電力と故障ノードの分布 (文献 [4]) より

ク内のノードでは故障が互いに関連する傾向が見られ、電源品質などの共通の環境要因が影響している可能性が示唆されている。また、文献 [4] 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と報告された。

## 4 研究目的

先行研究 [1] では、ノードの使用状況と故障の関係も調査されている。ただし、同研究における使用状況の分析は、平均ノード使用率やノードのジョブの数といった、比較的単純な指標に基づいている。さらに、同研究は、ハードウェア故障や環境要因など、HPC システム全体の故障傾向を広く分析している。さらに、文献 [4] では、電力ばらつきとノード故障との間に明確な関係が見られなかったという報告がある。本研究ではこの結果を踏まえ、ノードの故障率についてより詳細に考察する。ノードの故障率には、ノードの使用率だけでなく、ノードで実行されるジョブの特性が影響を与える可能性があると考えられる。本研究では、ジョブの特性に着目し、ノード故障にどのような影響を与えるかを詳細に分析することを目的とする。富岳の故障ログだけでなく、ジョブログ等の他のログと併せて分析することで、ノード故障の傾向や要因を明らかにすることを目指す。

## 5 今後の予定

- 富岳の故障ノードのログを分析し、故障の傾向や要因を明らかにする。
- 故障ノードのログを基に、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。

## 参考文献

- [1] El-Sayed, N. and Schroeder, B.: Reading between the lines of failure logs: Understanding how HPC systems fail, *2013 43rd Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)* (2013).
- [2] Fumiyoshi Shoji1, Shuji Matsui, M. O. F. S. T. T. A. U. K. Y.: Long term failure analysis of 10 petascale super-computer (2015).
- [3] TOP500: HPCG - November 2024, TOP500.org (online), available from<<https://www.top500.org/lists/hpcg/>

2024/11/> (accessed 2025-05-12).

- [4] 草場智也：A64FX スーパーコンピュータにおける電力ばらつきとシステム運用への応用，Vol. 39, pp. 1–39 (2025).
- [5] 理化学研究所計算科学研究センター：「富岳」について，(オンライン)，入手先<<https://www.r-ccs.riken.jp/fugaku/>> (参照 2025-05-12).