

令和 6 年 6 月 17 日現在

機関番号：32657

研究種目：基盤研究(B) (一般)

研究期間：2021～2023

課題番号：21H03449

研究課題名(和文) 中・長期障害発生予測に基づくシステム高信頼化技術の開拓

研究課題名(英文) Development of system reliability improvement technology based on medium- to long-term failure prediction

研究代表者

江川 隆輔 (Egawa, Ryusuke)

東京電機大学・工学部・教授

研究者番号：80374990

交付決定額(研究期間全体)：(直接経費) 12,900,000円

研究成果の概要(和文)：大規模化、複雑化が進む高性能計算システムにおける信頼性の向上、堅持のための要素技術に関する研究に取り組み、システムのログメッセージやヘルスマニタリング情報を用いた障害発生予測機構を開発した。また、高性能計算システムの高効率利用を目的に、システムの挙動を再現可能なジョブスケジューリングシミュレータを開発し、低消費電力ジョブスケジューリングアルゴリズムや緊急ジョブのためのジョブスケジューリングアルゴリズムを開発し、その有用性を示した。

研究成果の学術的意義や社会的意義

高性能計算システムは、科学技術やものづくりのための計算基盤としてばかりでなく、近年、気象予想技術、津波浸水被害予測技術など社会基盤として重要な役割を担っている。このため、システムを安定的、かつ効率的に利用することが強く求められている一方で、システムの高性能化に伴い、システムは大規模化複雑化の一途を辿り、システムの信頼性の向上、堅持が強く求められている。信頼性だけでなく、システムの効率的利用にも取り組んだ本研究は、将来の高性能計算基盤の運用の効率化に向けた基礎的研究と言えるものであり、社会的にも意味がある。

研究成果の概要(英文)：We have conducted research on elemental technologies to improve and maintain the reliability of high-performance computing systems, which are becoming increasingly large and complex. We developed technologies for collecting and aggregating system log messages and health monitoring information, and created a mechanism to predict failures using these data. Besides, to enhance the efficient use of HPC systems, we developed a job scheduling simulator capable of replicating system behavior and designed low-power job scheduling algorithms as well as algorithms for urgent jobs, demonstrating their effectiveness. These technologies have the potential to enhance the reliability and throughput of future HPC systems.

研究分野：計算機システム

キーワード：高性能計算 ジョブスケジューリング 障害 予測 計算システム

## 1. 研究開始当初の背景

スーパーコンピュータに代表される高性能計算システムは、気象予測、地震シミュレーション、創薬、気候変動のモデル化、天文学的観測データの解析など、非常に大規模かつ長い時間を要する計算を処理します。これらの計算の多くは我々の生活にも影響を及ぼし、安定して計算処理を行うことが可能なシステムの信頼性の維持、向上が強く求められている。さらに、計算中のシステム障害は、再計算やデータの修復に多大なコストを必要とするばかりでなく、システムのスループットを低下させる。一方で、近年のシステムの高性能化に伴い、高性能計算システムは大規模化、複雑化の一途を辿り、将来の高性能計算システムの平均故障間隔(MTBF)は大幅に短縮することが予想されており、長時間に及ぶアプリケーション実行を担保するためにも、高性能計算システムの信頼性、耐障害性の堅持は重要な課題となっている。

来たるエクス時代、ポストエクス時代の高性能計算システムにおいて、アプリケーション実行の信頼性を維持・向上させる取り組みとして、チェックポイントリスタート(以下 CPR)機構の開発が精力的に行われているが、これらは「障害発生後の被害を最小限にとどめる施策」と言わざるを得ない。さらには、システムの大規模化・複雑化に伴い、ジョブのスナップショット(リスタートするために必要なデータ)取得のための消費電力、時間、スナップショット容量の著しい増加、性能に与える影響が健在化しつつあり、MTBFの短縮に応じた頻繁なスナップショットの取得は現実的ではないことが知られている。将来の高性能計算システムにおいて、アプリケーション実行の信頼性を担保するためには、システム障害によるアプリケーションの停止を我々の日々の暮らしにおいて常に起こりえる「災害」と捉え、「防災・減災」の概念を取り入れることが必要だと考えられる。将来に起こりえる障害を発生前に予測し、障害発生が予想される計算リソースへのジョブのアロケーションを抑制するなど、事前に対応することで障害発生を抑制(防災)したうえで、障害を回避できなかった場合には必要最低限のスナップショットを取得する(減災)ことで、CPRのオーバーヘッドを軽減しながら、速やかにアプリケーションの実行を再開することが可能だと考えられる。

また、現在の大規模システムでは多数のモニタリング情報が採取可能になっている一方で、取得された情報と障害発生の因果関係が明らかになっていない。これらの因果関係を明らかにし、従来の対処療法的な高信頼化技術のみに頼らず、数時間から数日後の未来に起こりえる障害を事前に予測・検知し、信頼性を維持する機構の実現に向けた要素技術に関する研究は積極的に行われているとは言い難い。将来の大規模システムの平均故障間隔は数分になることが予想されている中、現在、提案・実用化されているプロアクティブな障害への対策は、システムのリアルタイムモニタリングのみに基づいており、高々数秒から数十秒先の障害発生予測に留まっている。このため障害に対する施策は減災志向な対応策に限られているのが現状である。

## 2. 研究の目的

本申請課題の目的は、従来の大規模高性能計算システムの高信頼化技術と大規模データ解析技術を融合させることで、将来起こりえる障害に対する対応策を導き出すことが可能な高性能計算システムの信頼化技術を確立することにある。

## 3. 研究の方法

本研究では、システムの高信頼化に向けて、障害発生予測、障害回避・復旧支援の2つの要素技術に着目し、課題を設定する。各課題とその関連技術が有機的、且つ密に連携しながら研究を進める。

### ・ 障害発生予測技術

大規模データ解析技術を用いて、実システムの運用を通して蓄積されたヘルスマニタリング情報と障害発生の因果関係を明らかにし、将来の高性能計算システムにおける中長期障害発生予測技術を確立すると共に、従来の短期障害発生予測技術との併用により予測精度の向上を図ることで、データ解析技術を用いた障害発生予測の実用性とその限界を明らかにする。

### ・ 障害回避・障害復旧支援技術

障害の発生が予測される計算資源へのジョブを割り当て回避しながら、高い性能と省エネルギーを担保可能な要件を明らかにし、アプリケーション実行の信頼性を高める運用技術ジョブスケジューリング技術の開発に取り組む。また、長期障害発生予測に基づくシステム保守を視野に入れ、将来の高性能計算システムのための高信頼・高効率運用指針に関する研究に取り組む。また、障害発生予測に基づくジョブスケジューリングだけでは防ぎきれない障害に対処するべく、アプリケーション実行の信頼性を担保するための「必要最低限」かつ速やかにジョブの再実行を可能と技術の研究開発に

取り組み, ジョブ障害発生予測, システム, 実行されるジョブの特性に応じて, システムのスループットとエネルギー効率を最大化する技術に関する研究に取り組む.

#### 4. 研究成果

障害発生予測技術においては, HPC システムや大規模なデータセンタなどにおいて, 現状では異常データが少ないことから教師あり学習による機械学習を用いての異常検知は困難であることを考慮し, 教師なし異常検知に分類されるオートエンコーダと呼ばれるニューラルネットワークモデルに焦点を当てた研究を進めた.

一般的な全結合オートエンコーダ(ディープオートエンコーダ・スパースオートエンコーダ)を用いての異常検知では得られる精度が限定的であることを示し, 精度向上の改善案として LSTM を組み込んだ LSTM オートエンコーダを用いた異常検知手法を提案した. 評価にはローレンス・リバモア国立研究所(LLNL) が運用していた Blue Gene/L の 215 日分のシステムログを用い[1], 畳み込みオートエンコーダ, GRU オートエンコーダとの比較を行い, 表 1 に示すように提案する LSTM オートエンコーダによる障害発生予測手法が最も高い精度と F1 スコアを達成可能であることを示した.

表 1 障害発生予測評価結果.

	Deep	Sparse	CNN	LSTM	GRU
Accuracy	0.999974	0.999976	0.999982	0.999988	0.999984
Precision	0.837838	0.849315	0.875	0.925373	0.898551
Recall	0.984127	0.984127	1	0.984127	0.984127
F1 score	0.905109	0.911765	0.933333	0.953846	0.939394

また, 提案手法と再帰的教師なし学習の異常検知手法である Recurrent Unsupervised Anomaly Detection(以下 RUAD)[2]との比較を行った. イタリア CINECA が公開している Tier-0 HPC システムである Marconi100 のヘルスマonitoringデータを用いて評価を行った. 本データは EXAMON と呼ばれる高性能計算システムにおけるデータ収集フレーム[3]により集約されたセンサーデータ 88 項目とシステムアラートログを使用し, センサーデータは各時間の特徴量としてシステムアラートログは各時間において異常が発生しているか否かを示す異常ラベルとして用いた. 評価の結果, 図 1 に示すように提案手法は RUAD と同等の精度でシステムの異常を検知できることを確認した.

高性能計算システムにおける障害発生予測・異常検知においては, 今回検討した手法が一定の精度で異常を検知可能であることを明らかにした. しかし, 高性能計算システムの運用期間が限定的であること, また採取可能なモニタリングデータが飛躍的に増加している一方で障害・異常の発生率がそれほど高くないことが, この予測問題を困難なものにしていることが確認できた. また, 多くの高性能計算システムでは, そのデータの膨大さからシステムのヘルスマonitoring情報の包括的に管理しきれず, 即座にデータ解析を行うのが困難であることも確認できた. これらは将来の高性能計算システムの高信頼化に向けた課題であると考えられる.

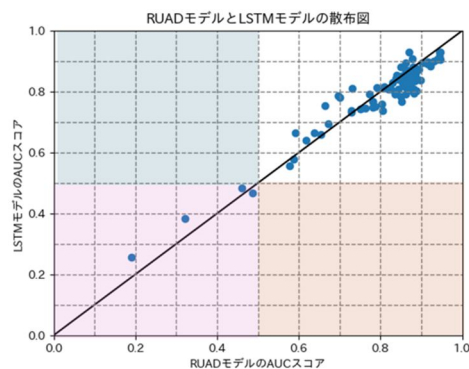


図 1 RUAD モデルと LSTM モデルの性能評価結果.

障害回避・障害復旧支援技術においては, 実システムのジョブ実行の挙動を再現可能なジョブスケジューリングシミュレータを開発した. シミュレータを用いてシステムのスループットの向上を目的としたジョブスケジューリングアルゴリズムや, システムのスループットを維持しながら緊急ジョブと呼ばれる優先度の高いジョブの実行を支援可能なスケジューリングの提案と評価を行い, その有用性を示している. また, ホスト CPU とアクセラレータから構成される高性能計算システムにおけるノード内のジョブの共有実行アルゴリズムの評価, 次世代高性能計算システムの性能評価にも取り組み, 将来の高性能計算システムにおける運用上の課題に関する検討も行った.

以上のように、本研究で得られた成果は将来の高性能計算システムの信頼性、スループット、運用コスト等を最適化するために有用な知見を与えるものであると考える。

[参考文献]

[1] Adam Oliner and Jon Stearley, “What supercomputers say: A study of five system logs,” In proceedings of 37th annual IEEE/IFIP international conference on dependable systems and networks (DSN '07), pp. 575-584. IEEE, 2007.

[2] Martin Molan, Andrea Borghesi, Daniele Cesarini, Luca Benini, Andrea Bartolini, RUAD: unsupervised anomaly detection in HPC systems. *Future Gener. Comput. Syst.*, 141, 542-554, 2022.

## 5. 主な発表論文等

〔雑誌論文〕 計11件（うち査読付論文 9件 / うち国際共著 3件 / うちオープンアクセス 1件）

1. 著者名 Riku Nunokawa, Yoichi Shimomura, Mulya Agung, Ryusuke Egawa & Hiroyuki Takizawa	4. 巻 1
2. 論文標題 Towards Conflict-Aware Workload Co-execution on SX-Aurora TSUBASA	5. 発行年 2022年
3. 雑誌名 Proceedings of Parallel and Distributed Computing, Applications and Technologies 2021	6. 最初と最後の頁 163-174
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-96772-7_16	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する
1. 著者名 Li Jiahao, Agung Mulya, Takizawa Hiroyuki	4. 巻 -
2. 論文標題 Evaluating the Performance and Conformance of a SYCL Implementation for SX-Aurora TSUBASA	5. 発行年 2022年
3. 雑誌名 Proceedings of The 22nd International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT ' 21)	6. 最初と最後の頁 36 ~ 47
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-96772-7_4	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する
1. 著者名 Ohmura Tatsuyoshi, Shimomura Yoichi, Egawa Ryusuke, Takizawa Hiroyuki	4. 巻 -
2. 論文標題 Toward Building a Digital Twin of Job Scheduling and Power Management on HPC System	5. 発行年 2023年
3. 雑誌名 Job Scheduling Strategies for Parallel Processing (JSSPP 2022)	6. 最初と最後の頁 47 ~ 67
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-031-22698-4_3	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Shimomura Yoichi, Musa Akihiro, Sato Yoshihiko, Konja Atsuhiko, Cui Guoqing, Aoyagi Rei, Takahashi Keichi, Takizawa Hiroyuki	4. 巻 -
2. 論文標題 A Real-time Flood Inundation Prediction on SX-Aurora TSUBASA	5. 発行年 2022年
3. 雑誌名 Proceedings of 29th International Conference on High Performance Computing, Data, and Analytics (HiPC)	6. 最初と最後の頁 192 - 197
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/HiPC56025.2022.00035	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Ide Kazuki, Takahashi Keichi, Shimomura Yoichi, Takizawa Hiroyuki	4. 巻 -
2. 論文標題 A Task-Parallel Runtime for Heterogeneous Multi-node Vector Systems	5. 発行年 2023年
3. 雑誌名 Proceedings of 23rd International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT 2022)	6. 最初と最後の頁 331 ~ 343
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-031-29927-8_26	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Sugawara Shunpei, Takahashi Keichi, Shimomura Yoichi, Egawa Ryusuke, Takizawa Hiroyuki	4. 巻 -
2. 論文標題 Equivalence Checking of Code Transformation by Numerical and Symbolic Approaches	5. 発行年 2023年
3. 雑誌名 Proceedings of 23rd International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT 2022)	6. 最初と最後の頁 373 ~ 386
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-031-29927-8_29	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Takizawa Hiroyuki, Sugawara Shunpei, Shimomura Yoichi, Takahashi Keichi, Egawa Ryusuke	4. 巻 -
2. 論文標題 Xevolver for Performance Tuning of C Programs	5. 発行年 2023年
3. 雑誌名 Sustained Simulation Performance 2021	6. 最初と最後の頁 85 ~ 93
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-031-18046-0_6	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Satake Moto, Takahashi Keichi, Shimomura Yoichi, Takizawa Hiroyuki	4. 巻 -
2. 論文標題 Balancing exploitation and exploration in parallel Bayesian optimization under computing resource constraint	5. 発行年 2023年
3. 雑誌名 Proceedings of The 18th International Workshop on Automatic Performance Tuning (iWAPT 2023)	6. 最初と最後の頁 706 - 713
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/IPDPSW59300.2023.00122	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Takahashi Keichi, Fujimoto Soya, Nagase Satoru, Isobe Yoko, Shimomura Yoichi, Egawa Ryusuke, Takizawa Hiroyuki	4. 巻 -
2. 論文標題 Performance Evaluation of Next-Generation SX-Aurora Tsubasa Vector Supercomputer	5. 発行年 2023年
3. 雑誌名 Proceedings of ISC High Performance 2023	6. 最初と最後の頁 359 ~ 378
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-031-32041-5_19	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Takizawa Hiroyuki, Takahashi Keichi, Shimomura Yoichi, Egawa Ryusuke, Oizumi Kenji, Ono Satoshi, Yamashita Takeshi, Saito Atsuko	4. 巻 -
2. 論文標題 AOBA: The Most Powerful Vector Supercomputer in the World	5. 発行年 2024年
3. 雑誌名 Sustained Simulation Performance 2022	6. 最初と最後の頁 71 ~ 81
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-031-41073-4_6	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Nunokawa Riku, Shimomura Yoichi, Agung Mulya, Egawa Ryusuke, Takizawa Hiroyuki	4. 巻 -
2. 論文標題 Conflict-aware workload co-execution on SX-aurora Tsubasa	5. 発行年 2023年
3. 雑誌名 CCF Transactions on High Performance Computing	6. 最初と最後の頁 1-14
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/s42514-023-00171-x	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

〔学会発表〕 計6件 (うち招待講演 1件 / うち国際学会 0件)

1. 発表者名 青柳嶺, 高橋慧智, 下村陽一, 滝沢寛之
2. 発表標題 計算特性に着目した実行時間予測に基づくリアルタイム洪水シミュレーションの動的資源割当
3. 学会等名 第185回HPC研究発表会
4. 発表年 2022年

1. 発表者名 石井翔, 高橋慧智, 下村陽一, 滝沢寛之
2. 発表標題 機械学習に基づくジョブスケジューリングのためのGANによるデータ拡張
3. 学会等名 第185回HPC研究発表会
4. 発表年 2022年

1. 発表者名 中井大貴, 大村竜義, 高橋慧智, 下村陽一, 滝沢寛之
2. 発表標題 緊急ジョブの優先実行を考慮したジョブスケジューリングに関する一検討
3. 学会等名 第187回HPC研究発表会
4. 発表年 2023年

1. 発表者名 江川隆輔
2. 発表標題 使える高性能計算機システムの実現にむけて
3. 学会等名 学術情報メディアセンターセミナー「時代に合ったHPCの活用」(招待講演)
4. 発表年 2023年

1. 発表者名 佐竹望都, 高橋慧智, 下村陽一, 滝沢寛之
2. 発表標題 探索と活用の調整による並列ベイズ最適化の効率化
3. 学会等名 第188回HPC研究発表会
4. 発表年 2023年



1. 発表者名 幸田 涼詩, 高橋 慧智, 下村 陽一, 滝沢 寛之
2. 発表標題 ベクトルプロセスを用いた統計的機械学習に関する研究
3. 学会等名 xSIG 2023
4. 発表年 2023年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	滝沢 寛之  (Takizawa Hiroyuki)  (70323996)	東北大学・サイバーサイエンスセンター・教授   (11301)	
研究分担者	谷村 勇輔  (Tanimura Yusuke)  (80415710)	国立研究開発法人産業技術総合研究所・情報・人間工学領域・主任研究員   (82626)	
研究分担者	滝澤 真一郎  (Takizawa Shinichiro)  (80550483)	国立研究開発法人産業技術総合研究所・情報・人間工学領域・主任研究員   (82626)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------