

平成 30 年 6 月 15 日現在

機関番号：15501

研究種目：基盤研究(C) (一般)

研究期間：2015～2017

課題番号：15K00171

研究課題名(和文) 動的システム再構成に基づくディペンダブルな並列ボランティア計算基盤

研究課題名(英文) Dependable and parallel volunteer computing systems based on a dynamic system reconfiguration

研究代表者

福士 将 (Fukushi, Masaru)

山口大学・大学院創成科学研究科・准教授

研究者番号：50345659

交付決定額(研究期間全体)：(直接経費) 3,500,000円

研究成果の概要(和文)：インターネットに接続されるパソコンなどの計算資源(ノード)を用いるボランティアコンピューティング(VC)において、ノード間で通信を伴う並列計算を実行可能にするディペンダブルな並列VCの実現方式に関する研究を行った。その実現には、新たな課題となる「計算ジョブを投入するノード群の選択」と「通信相手ノードの離脱への対応」を解決しなければならない。本研究では、VCで不可欠な冗長計算に着目し、これを応用して動的にシステム(計算ノード群)を再構成する方式により、両課題の同時解決を図った。本方式を用いた並列VCのプロトタイプを開発し、ディペンダブルな並列VCが実行可能であることを実証した。

研究成果の概要(英文)：Toward the realization of parallel volunteer computing (VC) using personal computers (i.e. nodes) on the Internet, this research project studied the system configuration method. For realizing dependable parallel VC, we need to solve the two challenging problem: how to configure node groups for job assignment and how to cope with node defection for inter-node communication. We focus on the redundant computation which is widely employed in the current VC, and solve the two problems simultaneously by the approach of dynamically reconfiguring node groups. We implemented the prototype parallel VC system based on the approach and demonstrated the feasibility of dependable parallel VC.

研究分野：並列分散システム

キーワード：ボランティア・コンピューティング ディペンダブル・コンピューティング ハイパフォーマンス・コンピューティング 並列計算 情報システム

1. 研究開始当初の背景

ボランティアコンピューティング(VC)は、インターネットに接続されるパソコンなどの計算資源(ノード)を一般のユーザから無償で提供してもらい、大規模な並列分散環境を安価に構築する手法である。組織内の信頼できるノードを用いるグリッドなどと比較すると、機器の準備や管理にコストがかからず、多数のユーザの貢献により高性能化が容易に実現できるという利点がある。

VCの研究は国外で先行しており、約25万ユーザによりスパコン並みの性能を達成した例[1]があり、難病に対する新薬の開発[1]や惑星パルサーの発見[2]など、実社会に貢献している。しかしながら、現行のVCでは、ノード間で通信を必要としないサーバ・クライアント型の分散計算のみが対象であり、通信を含む一般的な並列計算は実行できないという大きな制限がある。

一方で、一般的な並列計算は、グリッドなどの研究分野で扱われており、ミドルウェアとしてGlobus ToolkitのMPITCH-G[3]や、国内で開発されたNinf-C/Ninf-G[4]などがある。これらの研究では、システムとして確実な動作を重視するために、安定して動作する管理されたノードを対象とするものが多く、VCのように、インターネットに接続されるノードへの適用は困難である。

2. 研究の目的

本研究では、通信を含む大規模な並列計算をVCで実行可能にすることを目的とする。そのためには、計算ジョブを導入するノード群の選択、計算結果の信頼性の保証、通信ノード相手の離脱の各問題に対処することが必要になる。本研究では、冗長計算を基本とした動的システム再構成に基づくディペンダブルな計算基盤を研究する。プロトタイプのVCシステムの実装を行い、実アプリケーションを用いた性能評価を行うことで、本アプローチの実現可能性を実証する。本研究で明らかにする内容は、以下の通りである。

- (1) ディペンダビリティ要求(信頼度、ノード数、完了率など)を満たすノード群選択手法
- (2) 並列VC基盤の実装方法とそのオーバーヘッド
- (3) 本手法の実アプリケーションへの適用効果

3. 研究の方法

(1) ディペンダビリティ要求を満たすノード群選択手法

申請者らが研究してきた信頼度評価に基づく多数決法をもとに、ジョブの信頼度評価式を拡張し、ノードの信頼度に加えて、離脱率(起動からの経過時間や過去の稼働履歴から算出)、必要ノード数、計算の冗長度などを統合して考慮可能なノード群の選択問題へと帰着させる。評価式によりジョブの期待

信頼度や多数決成功率を推定し、計算誤りやノードの離脱が発生しても多数決を成立させることが可能なノード群の選択手法を開発する。計算機シミュレーションにより、VCの完了時間や多数決成功率などの性能を評価する。

(2) 並列VC基盤の実装方法とそのオーバーヘッド

開発したノード群選択法をもとに、ディペンダブルな並列VC基盤の実装方法を研究する。解決すべき各問題に対して、冗長計算を基本とした動的システム再構成のアプローチをとる。このために、参加ノード側でノードの信頼度と離脱率を反映させたクラスタリングを行い、ジョブのディペンダビリティ要求に応じて、それを処理するクラスタを生成し、動的に維持する手法を開発する。本手法を実装するために、Webサーバと連携したVCサーバとシステム常駐型のクライアントプログラムを開発し、クラスタ維持や再構成指示のためのメッセージオーバーヘッドなど、実環境における性能評価を行う。

また、ノード間通信が不可能な場合も想定し、ノード間通信をサーバ経由の通信に置き換える並列VCの実現方式も調査する。さらに、より多くの参加者を集めるために、Android端末に対応したクライアントプログラムを開発し、ディペンダブルなクライアントを実現するために、セキュリティ攻撃の検知手法に関する検討も行う。

(3) 実アプリケーションへの適用効果

画像処理分野の3次元形状復元処理や動画画像からの移動物体検出処理を例に、本手法の実アプリケーションへの適用効果を検証する。そのために、まずは、両処理の並列処理手法を考案し、処理の完了時間や通信時間に関する性能評価を行う。次に、(2)で開発した並列VC基盤を用いて、実性能を解明する。

4. 研究成果

(1) ディペンダビリティ要求を満たすノード群の選択手法の開発

各ノードの信頼度と離脱率を考慮しながらジョブの完了確率を推定するアプローチにより、要求を満たすノード群を選択可能にする手法を考案した。実際のVCで得られたトレースデータからノードの挙動をモデル化し、離脱率を推定した。

図1に、並列計算の並列度Dに対する計算完了時間の関係を示す(全ジョブ数500,全ノード数3,500)。ノード群のサイズが固定されている単純な手法と比較して、提案手法を用いることにより、並列計算の完了時間を約60%削減可能であることが分かった。これは、ノードの離脱が発生した場合でも、並列計算の停止をなるべく回避し、再実行を抑制できるためである。

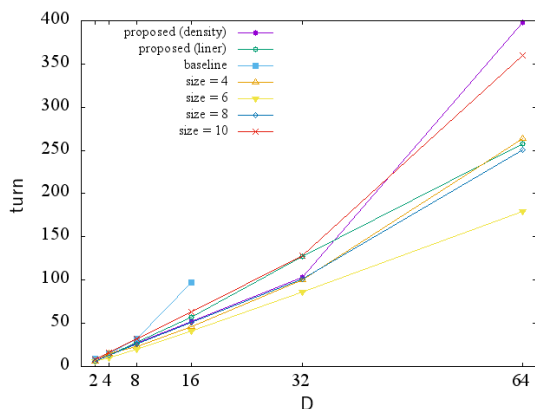


図1 並列度 D に対する実行時間(ターン)

(2) 並列 VC 基盤の実装方法とそのオーバーヘッド

ノード群の動的再構成手法

(1)で開発した手法をもとに、並列計算の各ジョブと通信グループの両方を冗長化するシステム構成を提案し、それらを動的に再構成する手法を提案した。本手法は、ノード間での局所的な情報の交換により、ノードの離脱が発生した場合でも、通信相手ノードの不在による並列計算の停止を抑制するものである。本手法をプログラム実装し、ノードの離脱が発生しても、他のノードが通信を代替し、システムを動的に再構成可能なことを確認した。また、再構成に要する時間(オーバーヘッド)は、並列度が 32 の場合、25 ミリ秒以下であることを明らかにした。

並列 VC システムのプロトタイプの開発

で提案した動的再構成手法をもとに、並列 VC のプロトタイプの開発を行った。そのために、Web で広く用いられている HTTP プロトコルで通信を行う VC サーバとクライアントの両方の開発を行った。VC サーバは、現在広く用いられている WEB サーバソフトウェアである Apache に、VC で必要になる基本機能をモジュールとして組み込み、内部の処理の待ち時間を削減するために、各処理をマルチスレッドで並行化する構成とした。クライアントにおいても、動的再構成法や VC 本来の計算処理、サーバとの通信を円滑に行えるように、各処理をマルチスレッドで並行化する構成とした。これらのプロトタイプを用いて、実環境においても、ノード群の再構成が正しく動作することを確認した。

図 2 に、実環境において、再構成時間を計測した結果を示す。実環境では、ノード間の通信遅延が発生するため、シミュレーションによる理論性能よりも再構成時間は長くなるが、並列度が 64 の場合でも、再構成時間は 2 秒以下であり、実用に耐えうる十分に小さな時間であることを明らかにした。

動的再構成を行うことで、通信相手ノードが離脱した場合には、他のノードが代理で通信に対処できるようになる。その場合、本来属する通信グループと、代理先の通信グループ

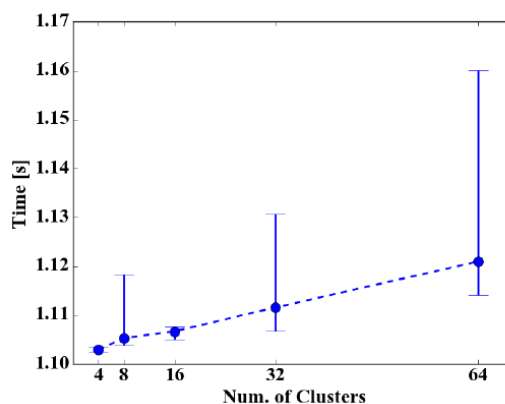


図2 並列度に対する再構成時間

での計算の進捗に差異がある場合には、いかに適切なデータを通信するかを考える必要がある。本研究では、通信データのログを保存しておくことにより、リクエストされた適切なデータを送信可能にしている。本機能を実装した通信ライブラリを用いて並列 VC の評価実験を行い、通信の代理が発生した場合でも、正しいデータを通信可能であることを確認した。

サーバ支援型の並列 VC

各ノードが属する組織などの取り決めにより、使用可能な通信ポートが制限される場合には、動的再構成や並列計算のためのノード間通信が実行できない可能性がある。そのような場合にも対処するために、ノード間の通信をサーバ経由の通信に代替するサーバ支援型の通信を実現するソフトウェアの開発も行った。並列通信ライブラリ MPI を用いてノード間で直接通信を行う PC クラスタと比較して、100 万個のデータのブロードキャスト時間は約 8 倍になったが、その完了時間は約 2 秒程度であった。前述したノード群選択手法と併用することで、ディペンダブルな並列 VC の実現手法の一つとして候補になり得る方式であることを確認した。

Android 端末に対応したクライアントアプリケーションの開発

VC で計算性能を向上させるには、より多くの参加ノードを集める必要がある。年々台数が増加しているスマートフォンによる VC の参加を目的に、Android 端末向けの VC クライアントの開発も行った。本クライアントは、VC サーバとの通信や計算ジョブの実行、CPU 使用率の制御などの機能を持つ Android アプリケーションとして開発した。前述した Web ベースの VC サーバを用いて VC の動作検証を行い、スマートフォン上で VC が動作可能であることを確認した。

ノードのセキュリティ攻撃検知

ディペンダブルな VC システムを構成するためには、インターネットで発生するセキュリティ攻撃も考慮する必要がある。本研究で

は、各ノードが自身に対するサービス妨害 (DoS) 攻撃を検知する検知手法を考案した。通常、各ノードは正確な時刻を保つために、時刻同期サービスを稼働させている。本手法では、被攻撃時に、時刻同期の時刻の調整量に揺らぎが発生することに着目して攻撃検知を行う。本手法をプログラム実装した検知モジュールを開発し、仮想的な DoS 攻撃や踏み台攻撃を行った。その結果、約 97% の検出精度を達成することを明らかにした。また、検知に必要なメモリ量も約 11Mbyte と非常に軽量であり、スマートフォンなどの小型のデバイスにも有用な手法であることを明らかにした。

(3) 実アプリケーションへの適用効果

VC の実アプリケーションとして画像処理分野の三次元形状復元処理や動画からの移動物体検出処理を対象に、並列化手法を検討した。いずれの処理においても、計算で発生するデータアクセスパターンの偏りに着目して、データ分割を行うものである。前者においては、各画像と仮想中心ボクセルを結ぶ直線の傾きを考慮したデータ分割を行い、単純なデータ分割手法と比較して、並列度が 32 の場合に、計算時間を約 60% 削減可能であることを明らかにした。後者においては、物体の輪郭に沿うような小領域でデータ分割を行い、並列度が 16 の場合に、逐次処理した場合と比べて、計算時間を約 72% 削減可能なことを明らかにした。

今後は、より大規模な並列 VC 環境で評価を行う予定である。

<引用文献>

- [1] Folding@home. <http://folding.stanford.edu/>
- [2] B. Knispel et al., "Pulsar Discovery by Global Volunteer Computing," *Science*, vol. 329, no. 5994, 2010.
- [3] I. Forester et al., "Globus: A Meta-computing Infrastructure Toolkit," *J. of HPCA*, vol. 11, no. 2, pp. 115-128, 2007.
- [4] Ninf-C/Ninf-G. <http://ninf.apgrid.org/>

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 11 件)

福土 将 (5 名, 2 番目), "IoT へのサイバー攻撃に向けた時刻同期サービスを利用したイベント検知手法", *電子情報通信学会論文誌, Vol. J101-D, No. 5*, pp. 742-753, 2018. (査読有)

DOI: 10.14923/transinfj.2017JDP7019

Y. Kambara, M. Fukushi, and T. Abe, "Parallelization Methods for Moving Objects Extraction Based on SLIC," *Proc.*

of The Twenty-third Int'l Symp. on Artificial Life and Robotics, 巻数無, pp. 474-477, 2018. (査読有)

S. Inohara and M. Fukushi, "Implementation and Evaluation of a Dependable Web-based Volunteer Computing System," *Proc. of The Twenty-Second Int'l Symp. on Artificial Life and Robotics*, 巻数無, pp. 530-534, 2017. (査読有)

M. Fukushi (5 名, 2 番目), "An NTP-based Detection Module for DDoS Attacks on IoT," *Proc. of IEEE Int'l Conf. on Consumer Electronics-Taiwan*, 巻数無, pp. 15-16, 2017. (査読有)

DOI: 10.1109/ICCE-China.2017.7990972

T. Sakai and M. Fukushi, "A Reliable Volunteer Computing System with Credibility-based Voting," *Journal of Information Processing*, Vol. 24, No. 2, pp. 266-274, 2016. (査読有)

DOI: 10.2197/ipsjjip.24.266

M. Fukushi, Y. Tateishi, and T. Abe, "A Parallel Method for Voxel-based Multiview Stereo," *IEEEJ Transactions on Image Electronics and Visual Computing*, Vol. 4, No. 2, pp. 136-144, 2016. (査読有)

K. Ochi and M. Fukushi, "A Group-Based Job Scheduling Method for Parallel Volunteer Computing," *Proc. of Third Int'l Symp. on Computing and Networking*, 巻数無, pp. 571-575, 2015. (査読有)

DOI: 10.1109/CANDAR.2015.74

S. Yasuda, Y. Nogami, and M. Fukushi, "A Dynamic Job Scheduling Method for Reliable and High-Performance Volunteer Computing," *Proc. of IEEE Int'l Conf. on Information Science and Security*, 巻数無, pp. 1-4, 2015. (査読有)

DOI: 10.1109/ICISSEC.2015.7370964

S. Tani, Y. Nogami, and M. Fukushi, "An Implementation of Credibility-based Job Scheduling Method in Volunteer Computing Systems," *Proc. of IEEE Int'l Conf. on Consumer Electronics-Taiwan*, 巻数無, pp. 37-38, 2015. (査読有)

DOI: 10.1109/ICCE-TW.2015.7216948

M. Fukushi (4 名, 3 番目) "A Performance Evaluation of Web-based Volunteer Computing using Applications with GMP," *Proc. of IEEE Int'l Conf. on Consumer Electronics-Taiwan*, 巻数無, pp. 41-42, 2015. (査読有)

DOI: 10.1109/ICCE-TW.2015.7216970

M. Fukushi (5名, 4番目), "Dynamic Job Scheduling Method based on Expected Probability of Completion of Voting in Volunteer Computing," IEICE Transactions on Information and Systems, Vol.E98-D, No.12, pp. 2132-2140, 2015. (査読有)
DOI: 10.1587/transinf.2015PAP0027

[学会発表](計 10件)

岡崎 有加, 池辺 純一郎, 福土 将, "Android 用 VC ワークアプリケーションの開発," IEEE CE ソサイエティ西日本合同チャプタ主催 1月研究会, 2018.

神原 勇太, 福土 将, 阿部 亨, "SLIC を用いた移動物体抽出処理の並列化手法," 電子情報通信学会 第 63 回機能集積情報システム研究会, 2017.

堀井 健太郎, 福土 将, "ボランティアコンピューティングにおける並列計算の実現に関する一考察," 電子情報通信学会 第 63 回機能集積情報システム研究会, 2017.

池辺 純一郎, 福土 将, "並列 VC のためのシステムモデルとグルーピングアルゴリズム," 平成 29 年度電気・情報関連学会中国支部連合大会, 2017.

越智 郁, 福土 将, "並列 VC の実現に向けたワークの特性を考慮した冗長計算手法," 電子情報通信学会 アシユアランスシステム研究会, 2017.

越智 郁, 福土 将, "並列 VC の実現に向けたノードの参加予測に基づくジョブスケジューリング法," 電子情報通信学会 第 61 回機能集積情報システム研究会, pp.1-6, 2016 .

境 岳志, 福土 将, "VC システムにおける信頼度に基づくジョブスケジューリング法の性能評価," IEEE CE ソサイエティ西日本合同チャプタ主催 1月研究会, 2016 .

立石 雄大, 福土 将, 阿部 亨, "大規模な多眼ステレオ処理に対する並列処理手法," IEEE CE ソサイエティ西日本合同チャプタ主催 1月研究会, 2016 .

越智 郁, 福土 将, "並列 VC の実現に向けたワークの離脱を考慮したジョブスケジューリング," 平成 27 年度電気・情報関連学会中国支部連合大会, 2015 .

猪原 聡二, 福土 将, "高信頼 Web ベース VC の実現に向けたサーバの実装と性能評価," 平成 27 年度電気・情報関連学会中国支部連合大会, 2015 .

6. 研究組織

(1) 研究代表者

福土 将 (FUKUSHI Masaru)
山口大学・大学院創成科学研究科・准教授
研究者番号: 50345659

(2) 研究分担者

阿部 亨 (ABE Toru)
東北大学・サイバーサイエンスセンター・准教授
研究者番号: 80222652