

課題名 (タイトル) :

通信・ファイル I/O に関する研究開発

利用者氏名 :

○石川 裕\*、高木 将通\*、堀 敦史\*、Gerofi Balazs\*、森江 善之\*、畑中 正行\*、小倉 崇浩\*、Martsinkevich Tatiana\*、亀山 豊久\*

理研での所属研究室名 :

\*計算科学研究機構 フラッグシップ 2020 プロジェクト システムソフトウェア開発チーム

1. 本課題の研究の背景、目的、関係するプロジェクトとの関係

本課題では、ポスト「京」で搭載される機能を有する FX100 の通信・ファイル I/O 機能を用いたシステムソフトウェアのプロトタイプ実装および性能検証を実施することを目的とする。我々は、ポスト「京」向けの OS、MPI ライブラリ等のシステムソフトウェアの研究開発を行なっている。このようなシステムソフトウェアの開発は、現在存在しないターゲット計算機のハードウェア開発と並行して進める必要がある。このため、ポスト「京」と同様の機能を有する FX100 を用いる実装実験が重要となる。

2. 具体的な利用内容、計算方法

今回は、ポスト「京」に向けて MPI インターフェースの定義が予想される永続型集団通信のプロトタイプ実装およびその性能、コストに関する評価を実施した。実装対象としては「京」でも実装されたブロードキャストの Trinaryx3 アルゴリズムを用いた。

FX100 では、アシスタントコアおよび Tofu オフロード機能を備えている。そこで、ターゲットである永続型ブロードキャストに関して Tofu オフロード機能を用いた実装とプログレススレッドを用いた実装の 2 種類のプロトタイプ実装を行なった。

3. 結果

今回の評価では、永続型ブロードキャストの 2 種類のプロトタイプ実装に対して通信性能および永続型集団通信の開始関数のコストについて評価を行った。

Tofu オフロード機能を用いた実装では 1KiB から

64MiB までのメッセージサイズでプログレススレッドを用いた実装に対して低遅延であることがわかった。これは、通信コマンド列を一度に投入できる Tofu オフロード機能により CPU と Tofu 間のインタラクションが削減されたからだと考えられる。

永続型集団通信の開始関数の実行時間は、プログレススレッドを用いた実装および Tofu オフロード機能を用いた実装ともに短いものとなった。プログレススレッドを用いた実装では通信コマンド発行などは行わず、集団通信の開始制御のみのため、短い実行時間になったと考えられる。また、Tofu オフロード機能を用いた実装では集団通信開始時の処理が 1 度の通信コマンド列の投入のみとなり、実行時間が短いものとなった。

他に、Tofu オフロード機能を用いた実装では、この開始関数のコストが通信コマンド数に比例して増加することなど課題があることもわかった。

4. まとめ

Tofu オフロード機能を用いた場合とプログレススレッドを用いた場合の 2 種類の永続型ブロードキャストのプロトタイプ実装を行い、それぞれの通信性能や永続型集団通信の開始関数のコストについて評価を行った。これにより、各実装の特性など今後永続型集団通信の開発を行うにあたり有用な結果を得た。

5. 今後の計画・展望

まずは、今回至らなかったスケーラビリティの評価、計算と通信の隠蔽効果の評価等を実施する。また、今回はブロードキャストのみの実装となったが、他の永続型集団通信についても同様に実装・評価を行う。これらの実装・評価を元にポス

平成 29 年度 利用報告書

ト「京」向けの永続型集団通信の適切な実装を行う予定である。