

## 大規模 CFD 解析におけるポスト処理効率化のためのデータステージング技術に関する研究

報告書番号：R20JACA42

利用分野：JSS 大学共同利用

URL：<https://www.jss.jaxa.jp/ar/j2020/14486/>

### ● 責任者

高橋慧智, 奈良先端科学技術大学院大学

### ● 問い合わせ先

高橋慧智(keichi@is.naist.jp)

### ● メンバ

高橋 慧智

### ● 事業概要

CFD シミュレーションの大規模化にともない、従来のように全てのシミュレーション結果を並列ファイルシステムに保存し、シミュレーション終了後にポスト処理を実行することはストレージ容量や IO 帯域幅の制約により困難になると予想される。そのため、シミュレーション実行中にポスト処理プログラムへ計算結果をリアルタイムに転送する、データステージング技術が注目されている。本研究では、JSS2 に代表される HPC 環境におけるデータステージング技術の実現可能性を検討し、データステージングミドルウェア、および、HPC 環境に求められる要件を分析する。

### ● JAXA スーパーコンピュータを使用する理由と利点

JSS3 はアーキテクチャの異なる HPC システム (TOKI-SORA) と汎用システム (TOKI-RURI) から構成され、2 システム間で通信が可能であるため。

### ● 今年度の成果

今年度より利用可能になった JSS3 システム上で、米国 Oak Ridge 国立研究所にて開発されているデータステージング用ミドルウェア ADIOS2 の動作試験およびステージング通信の性能評価を実施した。SORA の計算ノード上で富士通コンパイラ 4.4.0 (clang モード) を用い ADIOS2 v2.7.1 のビルドを試みたところ、特に問題なく完了し全ユニットテストが成功した。なお trad モードでは CMake がコンパイラを正しく認識しないため、CMake 側で富士通コンパイラに対応する必要があると思われる。

次に ADIOS2 に含まれる iotest ユーティリティを用い、SORA 上でステージング通信の性能評価を行なった。iotest で同一プロセス数の 2 つのアプリケーションを模し、アプリケーション間で 1 プロセスあたり 100MB のデータを送受信させた。ADIOS2 は様々な通信バックエンドを提供しているが、

ここでは最も性能が高いと思われる MPI を用いたステージング通信エンジン InSituMPI と SSC (Strong Staging Coupler) を比較した。SSC エンジンについては、さらに TwoSided, OneSidedPostPush, OneSidedPostPull の3つのモードの性能を比較した。TwoSided は1対1通信関数, OneSidedPostPush/Pull は片側通信関数を用いて実装されている。

図1に2アプリケーション間の通信帯域幅を示す。まず2ノード間では、InSituMPI エンジンで5.5GB/s, SSC エンジンではいずれのモードでも5.9GB/sの帯域幅が得られた。Tofu-D 相互結合網のリンク帯域幅は6.8GB/sであるため、InSituMPI エンジンと SSC エンジンでそれぞれリンク帯域幅の81%および87%の実効帯域幅が得られた。ノード数がスケールアウトすると InSituMPI が SSC に比べ性能が高くなり、512ノードでは InSituMPI が418GB/s, SSC がモードに関わらず270GB/sであった。

以上の評価により、SORA システムの相互結合網を活用した高いステージング通信性能を実現できることが明らかになった。ただし、高並列時に SSC エンジンの性能が InSituMPI エンジンに比べ劣る原因が不明であるため、調査する必要がある。また、SORA/RURI の異種アーキテクチャのシステム間の通信についても検証を行う必要がある。

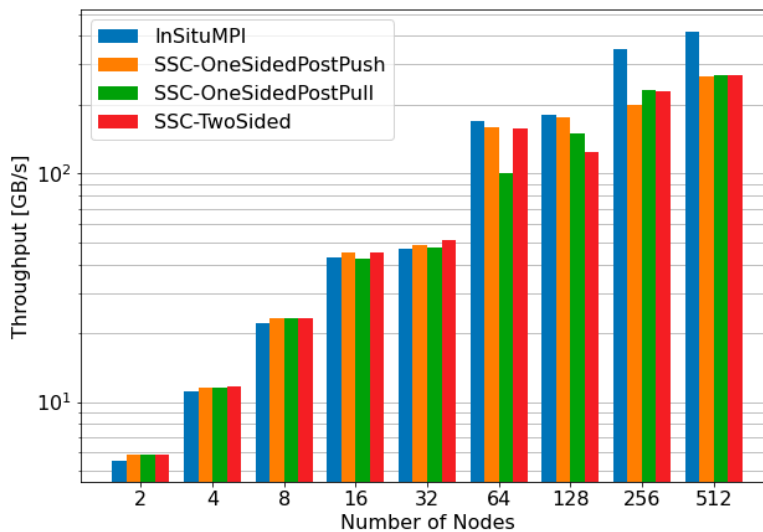


図1: adios\_iotest による2アプリケーション間のステージング通信性能 (48 プロセス/ノード, 5回計測の最大値)

● 成果の公表

なし

## ● JSS 利用状況

## ● 計算情報

プロセス並列手法	MPI
スレッド並列手法	非該当
プロセス並列数	1 - 24576
1 ケースあたりの経過時間	10 分

## ● JSS2 利用量

総資源に占める利用割合※1 (%) : 0.00

内訳

計算資源		
計算システム名	コア時間(コア・h)	資源の利用割合※2 (%)
SORA-MA	220.60	0.00
SORA-PP	0.00	0.00
SORA-LM	0.00	0.00
SORA-TPP	0.00	0.00

ファイルシステム資源		
ファイルシステム名	ストレージ割当量(GiB)	資源の利用割合※2 (%)
/home	9.54	0.01
/data	95.37	0.00
/tmp	1,953.13	0.17

アーカイバ資源		
アーカイバシステム名	利用量(TiB)	資源の利用割合※2 (%)
J-SPACE	0.00	0.00

※1 総資源に占める利用割合：3つの資源(計算,ファイルシステム,アーカイバ)の利用割合の加重平均

※2 資源の利用割合：対象資源一年間の総利用量に対する利用割合

## ● JSS3 利用量

総資源に占める利用割合※1 (%)： 0.01

内訳

計算資源		
計算システム名	コア時間(コア・h)	資源の利用割合※2 (%)
TOKI-SORA	26,332.37	0.01
TOKI-RURI	0.00	0.00
TOKI-TRURI	0.00	0.00

ファイルシステム資源		
ファイルシステム名	ストレージ割当量(GiB)	資源の利用割合※2 (%)
/home	9.54	0.01
/data	95.37	0.00
/ssd	95.37	0.05

アーカイバ資源		
アーカイバシステム名	利用量(TiB)	資源の利用割合※2 (%)
J-SPACE	0.00	0.00

※1 総資源に占める利用割合：3つの資源(計算,ファイルシステム,アーカイバ)の利用割合の加重平均

※2 資源の利用割合：対象資源一年間の総利用量に対する利用割合