

1. 将来のスパコン, 気象学=計算機科学連携

石川 裕*

2014年度からスタートした文部科学省フラッグシップ2020プロジェクトにおいて, 理化学研究所(以降, 理研)は次世代超高速電子計算機システム(以降, ポスト「京」)の開発・整備を担っている。ポスト「京」の開発では, 2020年までにシステムとアプリケーションのCo-designにより, 気象学を含む社会的・科学的課題の解決に貢献し, 世界トップレベルで幅広い課題に対応できる汎用システムを実現することが目標となっている。理研は公募により富士通を協働開発先に決め, また, フラッグシップ2020プロジェクトにおいて文部科学省が決めた9つの重点課題実施機関とも協働してポスト「京」の基本設計を行っている。

ポスト「京」は, 汎用プロセッサをコアに持つミニコアCPUを計算ノードに持ち「京」と同様の6次元メッシュ・トラスネットワークで接続される。9重点課題実施機関から提供されたアプリケーションコードをターゲットアプリケーションとしCPUアーキテクチャのパラメータ(メモリ階層, コア接続方式等)を決めている。

ファイルシステムは半導体デバイスで構成される第1階層ストレージ, ハードディスクで構成される第2階層ストレージ, アーカイブ目的の第3階層ストレージで構成される。ファイルI/Oの性能および機能要件をまとめるために, ポスト「京」においてターゲットアプリケーション群が利用するノード数に対して, どのようなパターンでファイルI/O処理をするのかヒアリングした。さらに, アプリケーションのファイルI/Oパターンを実行時に収集記録するツールを用いて「京」においてターゲットアプリケーションを実

行し, ヒアリング結果の確認を行った。また, 「京」におけるストレージの利用状況からポスト「京」の必要ストレージ容量を見積もった。これら要件から階層化ストレージの性能と容量を決めている。

科学技術振興機構CREST「『ビッグデータ同化』の技術革新の創出によるゲリラ豪雨予測の実証」(研究代表者:三好建正)では30秒サイクルで,

- i) 天気予測シミュレーションを同時に100ケースアンサンブルシミュレーションするジョブ群実行
- ii) これら結果と次世代型フェーズドアレイ気象レーダーで取得可能な30秒毎の3次元観測データを用いてデータ同化するジョブ実行
- iii) 30分後の予測シミュレーションジョブ実行

を行う。シミュレーションコードとデータ同化コードはそれぞれ独立して作られており, 現状はファイル渡しでデータが交換されている。30秒サイクルという実時間性を実現するためにはファイルI/O時間を短縮する必要がある。ファイルI/OのAPI(Application Program Interface)を変更することなくジョブ間通信によってデータを渡す方式の設計を行っている。また, 観測データはインターネット経由で送られる。観測設置場所で観測されたデータは観測設置機関のLANからインターネット経由で理研計算科学研究機構のアクセスポイントに送られ, そこからデータ同化しているジョブにデータを転送しなければならない。現在, このようなデータ転送の遅延時間を調査している。

* 理化学研究所計算科学研究機構。

yutaka.ishikawa@riken.jp

© 2017 日本気象学会