サイエンティフック・システム研究会　　2008 年 8 月 27 日
SS 研　HPC フォーラム 2008

次世代 HPC における大規模データの可視化環境

理化学研究所
小野謙二
keno@riken.jp

アブストラクト

超並列計算機が生み出す大規模なデータは，それを効率的に扱うこと自体がチャレンジングな課題であり，ポスト処理はこれまで以上に解析のボトルネックになる．可視化プロセスはシミュレータとは異なり，各ユーザ毎の多様な作業シナリオが存在する．このため，可視化システムは作業者の主観的な側面が強く現れ，機能性とともにユーザビリティが重要視される点が特徴的である．本講演では，大規模なデータを効率よく扱う上での問題点と要素技術について概要を述べるとともに，現在設計を進めている大規模可視化システムの開発について報告する．

キーワード

Large-scale data visualization, Knowledge discovery, simulation, framework, usability

講演概要

本講演では，理研で開発を進めている次世代スーパーコンピュータをはじめ，既に稼働している汎用的な大規模計算機システムでの利用を想定した大規模データ可視化システム，およびこの可視化システムを中心とした包括的なポスト処理システムのコンセプトについて述べる．基本となる考えは，研究者の「考えるプロセス」の支援である．先端的なシミュレーションは高性能な超並列分散環境で実現可能になる一方で，分散並列環境のポスト処理は必要な分析・可視化処理を行うための準備に多大な労力を強要する．この点は，現象の分析や理解を進める上で大きな弊害となる．また，シミュレーションのデータサイズの増大はポスト処理にかかる時間や作業効率・生産性に大きく影響する．さらに，先端的な計算機資源を用いた研究プロジェクトはもはや一人の研究者だけの研究ではなく，異なる分野・組織の研究者間のコラボレーションへ進展していくであろう．このような状況を鑑みると，先端研究プロジェクトでは，大規模データおよびそこから引き出された知識の共有という枠組みが重要な役割を担うことが考えられる．別の視点からは，可視化研究の振興，ソフトウェア資産の有効活用，あるいは継続的なメンテナンスによる研究支援という側面も重要になってくる．

これらの視点から組み立てたポスト処理システムのコンセプトについて，皆さんのご意見を伺い，更によりよいシステムへとブラッシュアップしたいと考えています．