

# グリッドデータファームによる大規模分散ストレージの構築とサイエンスクラウド技術の研究

村田 健史、渡邊英伸(独)情報通信研究機構)

目的: HPC/HTC(High Throughput Computing)を目的として、スパコンシミュレーションデータのポスト処理をNICTサイエンスクラウドで高速に行うための基礎的な実験を行った。

内容: スパコンによる3次元シミュレーションデータ(64ファイル・合計150GB)を、Gfarm/Pwrakeを用いて4台の計算機で可視化する。データは、I/O(Read)時間が可視化処理時間に比して無視できないため、I/O時間をどのように節約するかが重要である。

結果: Gfarm/Pwrakeにより、4台の計算機をGfarmのFSN(File System Node)とCN(Client Node)を兼務させる。①CNが自身がFSNとして管理するファイルのみを可視化処理する、②2台については①と同じとし、残り2台については異なるFSNが管理するデータファイルの処理を行うの2通りについて比較した。その結果、データ処理時間はほぼ一定であるがI/O時間が②の方が①よりも数倍必要となることが分かった。

利用した計算機: Gfarm計算機  
CPU時間: 20時間  
使用メモリ: 20 GB  
ベクトル化率: 95%  
並列化: 8並列

