

スーパーコンピューター結合とそれを活用した超大規模シミュレーション計算によるナノテクノロジー研究者用ネットワークの実現

一関京子^{A)}、川添良幸^{B)}、高橋 英明^{C)}、豊岡 雅人^{D)}、安達 斉^{D)}、遠藤 順一^{D)}

^{A)}東北大学金属材料研究所 技術室

^{B)}東北大学金属材料研究所計算材料学研究部門

^{C)}日本ヒューレットパッカード株式会社

^{D)}株式会社 日立東日本ソリューションズ

概要

本ナノテクノロジー研究者用ネットワークは、スーパーSINET を最大限に活用して、東北大学金属材料研究所(以下金研)、東京大学物性研究所(以下物性研)、岡崎国立共同研究機構分子科学研究所(以下分子研)、九州大学情報基盤センター(以下九大)および日本原子力研究所(以下原研)の所有する5台の遠隔スーパーコンピューターを結合し、分散処理による超大型材料設計シミュレーション計算を実行しようとするものである。平成14年度、金研で開発中の全電子混合基底法 TOMBO(TOhoku Mixed-basis Orbitals ab initio program)を使用して分子研と金研との間で実証実験を行い、平成15年度中には物性研との拠点間並列実行を行う予定であるので、それらについて報告する。

1 ナノテクノロジー研究者用ネットワーク

1.1 スーパーSINET

スーパーSINETは文部科学省国立情報学研究所(以下NII)が、平成14年1月4日より運用を開始した10ギガビットの光通信技術を用いる世界最高速の研究用の高速通信網で、高エネルギー・核融合科学研究、遺伝子情報解析研究(バイオインフォマティクス)、宇宙科学・天文学研究、ナノテクノロジー研究、スーパーコンピューター等を連動する分散コンピューティング(GRID)等の基盤用ネットワークとして用いられている。また平成15年度からは、原研等が開発しているITBL¹⁾(IT-Based Laboratory)計画についてもスーパーSINETが基盤整備しているため、ITBLが利用できるようになり、アプリケーション実行環境も整備されてきている。

1.2 ナノテクVPN

世界的にみてもほとんど試みられたことのないスーパーコンピューターのネットワーク経由結合による超大規模ナノテクノロジー用シミュレーション計算環境(以下、ナノテクVPN)をスーパーSINET上に実現するため、金研・物性研・分子研・九大の4機関によるVPN(Virtual Private Network)を構築し、平成14年10月1日に各機関間を1Gbpsで接続し効率測定が開始された。平成15年10月1日には北陸先端技術大学院大学、広島大学が1Gbpsで接続され、シミュレーション計算結果の可視化なども提案された。現在のネットワーク構成を図1に示す。

1) ITBL計画では計算機シミュレーションの活用により、理論や実験に代わる研究開発の手段を提供し、幅広い層での研究開発の効率化を目指し、異なる組織の計算機資源、ソフトウェア、データベース等をネットワーク上で共有化することにより共同研究を容易に実現できるシステム基盤を開発している。ITBL計画のとの連携研究の一環としてシステム基盤及び計算機を利用している。

ナノテクVPN構成概念図

平成15年10月開通

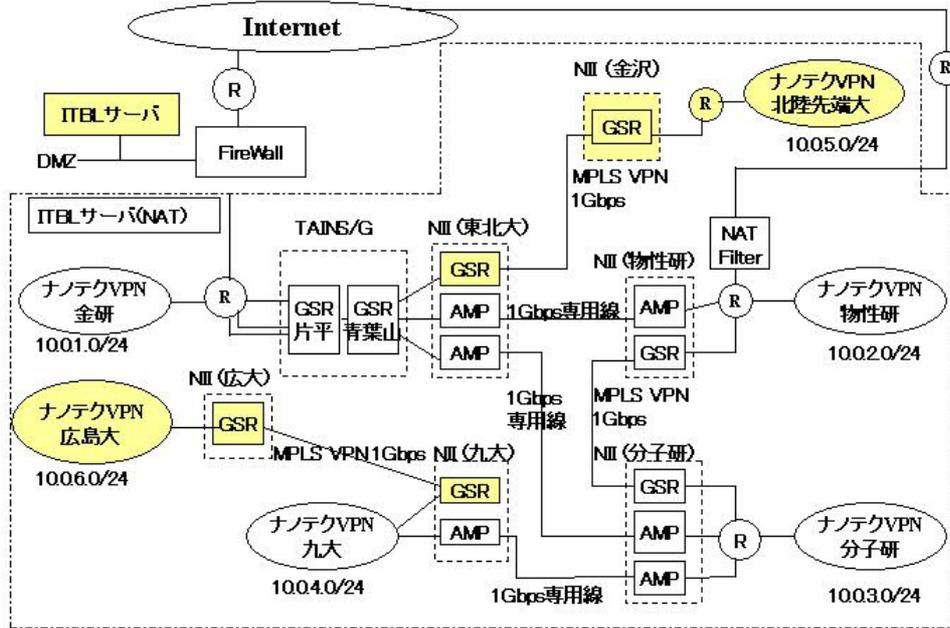


図 1. ナノテク VPN 構成概念図

1.3 ナノテクノロジー研究者用ネットワーク

平成 15 年 2 月に、金研のスーパーコンピューターSR8000(8 ノード)と分子研の SR8000(6 ノード)を、平成 15 年 12 月には物性研の SR8000 (4 ノード)をナノテク VPN に 1Gbps で接続し、他のトラフィックに邪魔されない環境を作成した。また、ITBL サーバーを DMZ (DeMilitarized Zone) と使用するスーパーコンピューターが接続されているネットワークに接続することにより、原研のスーパーコンピューターもナノテク VPN に接続されているものと同様に使用できるようになった。このナノテクノロジー研究者用ネットワーク環境を図 2 に示す。

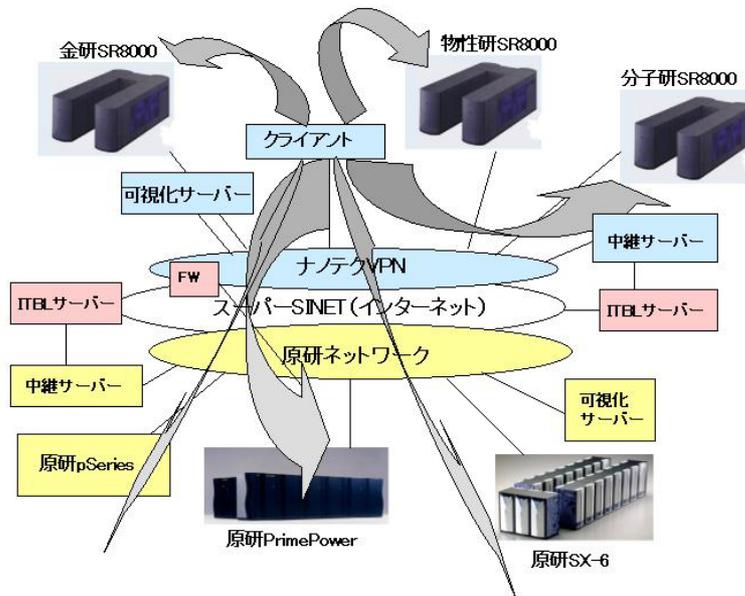


図 2. ナノテクノロジー研究者用ネットワーク

2 ナノテクノロジー研究者用ネットワークの実行環境

金研が開発した TOMBO プログラムの初期座標として異なる原子座標を入力データとして与え、数百件オーダーのジョブを同時並列実行してシミュレーション計算を行い、最適な原子構造の探索を行うパラメータサーベイ実行環境を平成 14 年度に開発した。平成 15 年度は、単一の計算機では実行できない大規模な材料設計シミュレーション計算ができる環境の開発を行っており、GW 近似を加えた全電子混合基底法第一原理計算プログラム(以下 GW プログラム)の拠点間並列実行を予定している。

2.1 パラメータサーベイ実行環境

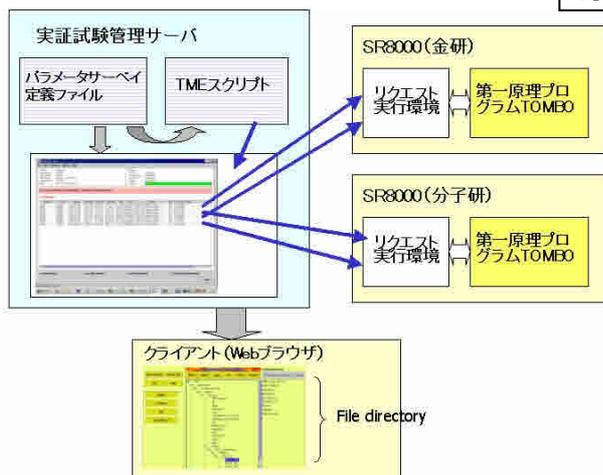


図 3. パラメータサーベイ実行環境

研究テーマ、アプリケーション名、ジョブグループ、開始時刻、終了時刻、経過時間、最良の計算結果

リクエスト数(トータル、未実行、実行中、終了数)およびバー表示

パラメータサーベイ状況表示
リクエスト番号、実行ホスト名、状態、開始/終了時刻、経過時間、ION/SCF ループ回数、エネルギーの収束状況、電荷密度の収束状況

図 4. パラメータサーベイサマリ

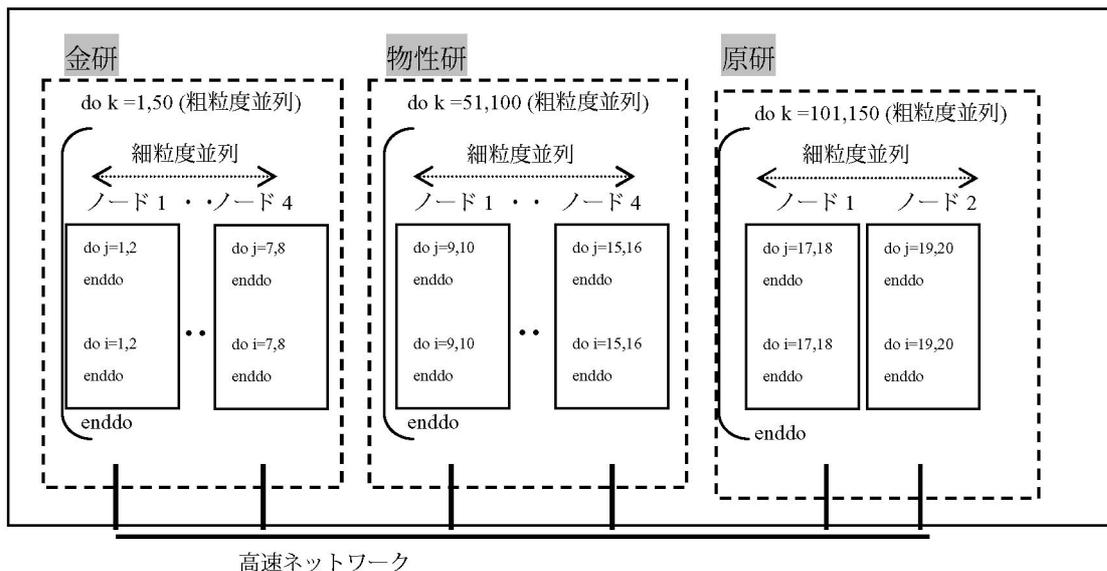
このパラメータサーベイ実行環境では、シミュレーション計算実行状況および結果を Web ブラウザ上にリアルタイムに表示できるので、エネルギー値、電荷密度値でソートすることが可能である。さらに実行結果を可視化することができる。この機能により、計算途中でチェックが可能となり、無駄なタスクを消去しながら最大効率で計算が実行できる。

図 5. 水クラスレートのパラメータサーベイサマリ画面

図 6. 水クラスレートの可視化結果表示画面

2.2 拠点間並列実行環境

GW プログラムは金研が開発した GW 近似による第一原理計算プログラムであり、局所密度近似を超えて、バンドギャップやイオン化ポテンシャルの絶対値算定を行なうことが可能である。そのために大量の計算量とメモリ量を必要とする。今回、この GW プログラムを拠点間にまたがって実行することにより、計算量とメモリ使用量の分散化を図り、これまで以上に大規模な計算を可能とするためのプログラム並列化を実施した。並列化手法には、粗粒度および細粒度の 2 種類の並列化を実施する。粗粒度並列化の目的は、大きな粒度でメモリ容量および演算量を各プロセスに分散化することである。細粒度並列化の目的は、粗粒度並列化した処理単位の内部について、細かい単位でプロセスに分散化することによりプロセスあたりの計算負荷を軽減することである。図 7 に今回予定している並列化 GW 計算の実行形態を示す。



分散化したプログラムによる並列実行

(粗粒度並列数 3、細粒度並列数 金研 4、物性研 4、原研 2 の場合の例)

図 7 . 拠点間並列実行形態

3 今後の計画

平成 15 年度は新たに、現在使用している金研開発の TOMBO、GW 以外のアプリケーションの開発を行っている広島大学および計算結果の可視化データのデータベース化を試みている北陸先端技術大学院大学がナノテク VPN に接続された。平成 16 年度は、それらの実用化を行うとともに、各拠点の機種異なるスーパーコンピュータそれぞれの特性を活かした連成計算を行い、仮想スーパーコンピューティングシステムとしてより充実したシステムにしていきたいと思っている。

【 謝辞 】

ナノテクノロジー研究者ネットワークの構築および実証実験において、分子研電子計算機センターの水谷文保技官、物性研物質設計評価施設の高山一施設長、吉本芳英助手には多大なご協力を得た。また、ITBL の環境設定および原研の計算機利用に関しては、原研計算科学技術推進センターの福田正大客員研究員、木村和幸研究員、大谷孝之課長代理に大変お世話になった。また、すべての基盤である高速ネットワークの敷設は、NII 開発・事業部の郷原正好ネットワークシステム課長補佐のご尽力がなければ構築できなかった。ここに記してすべての方々に深謝する。