

Quantum Chemistry GRID/Gaussian Portalの構築

西川 武志^{*,†} 長嶋 雲兵^{*,†} 関口 智嗣^{*,††}

Gaussian (<http://www.gaussian.com/>) は量子化学の専門/非専門を問わず化学者に広く使われる計算化学コードである。従って Gaussian ジョブは、殆どのスーパーコンピュータセンターにおいてキューに投入されたジョブの圧倒的多数を占めるが、幾分かは不適切な計算資源、すなわちキューや計算機に投入されたりしている。Gaussian ジョブが消費する CPU サイクルは入力パラメータに依存して際立って変化するので、ローカルな計算環境からの最も適切な計算のリソースを選択することはユーザーにとって困難である。我々はグリッド技術を適用することにより高速ネットワーク環境上に「Quantum Chemistry GRID/Gaussian Portal」を構築し、個々のシステム環境の仕様を知らずとも高価な計算資源を能率的に利用することを目指している。「Quantum Chemistry GRID/Gaussian Portal」は Web インターフェース、メタスケジューラー、計算資源、アーカイブ資源、グリッドインフラウェアから構成されている。

Building of Quantum Chemistry GRID/Gaussian Portal

TAKESHI NISHIKAWA^{*,†} UMPEI NAGASHIMA^{*,†}
and SATOSHI SEKIGUCHI^{*,††}

Gaussian (<http://www.gaussian.com/>) is a code widely used in computational chemistry research by quantum and non-quantum chemists. So, Gaussian jobs are the majority of the number of queued, but some are queued inadequate computational resources, queues or machines for the job at almost supercomputer center. Since consuming CPU cycles of Gaussian jobs vary significantly depending on the input parameter, it is difficult for users to chose the most adequate computational resources from local computing environment. By deploying grid technology on a top of high speed network environment, "Quantum Chemistry GRID/Gaussian Portal" attempt efficiently to utilize costly computational resources without knowing the specifications of each system environment. "Quantum Chemistry GRID/Gaussian Portal" consists of Web interface, Meta-scheduler, computational resources, archival resources, and Grid Infracwares.

1. はじめに

筆者らが属する産業技術総合研究所(産総研)グリッド研究センター²⁾は我が国におけるグリッド技術研究開発の中核拠点となることを目指し、最新のグリッドミドルウェア技術の開発や大規模高速計算システムの活用等によるテストベッド構築と実証システム

の開発を中心として、グリッド技術の飛躍的な高度化と体系化に貢献する研究開発を行うことを目標としている。その目標のうち大規模高速計算システム実証システム開発の応用例としてグリッド技術を適用し、量子化学計算プログラム Gaussian¹⁾のジョブを高速処理する Quantum Chemistry GRID/Gaussian Portal (QC GRID/Gaussian Portal) を構築したので報告する。同じく筆者らの属した/する産総研先端情報計算センター (TACC: Tsukuba Advanced Computing Center)³⁾では高性能計算機システムとして高性能計算機システムとして IBM RS6000/SP (128 ノード、256GB メモリ) を保有している。平成 12 年度の利用状況について分析を行ったところ、総 CPU 時間の内昨年度の月平均で 70%、最大 84% が Gaussian の実行に費やされていることが判明した。さらに TACC にライブラリ登録されたアプリケーションの総 CPU 時間の 93% が Gaussian によることも判明した。TACC では Gaussian を並列実行運用しておらず、HPS とい

* 現在、産業技術総合研究所 グリッド研究センター

Presently with Grid Technology Research Center, National Institute of Advanced Industrial Science and Technology, METI (AIST)

† 2002 年 4 月まで、先端情報研究センター

Tsukuba Advanced Computing Center (TACC)

†† 先端情報研究センター

Tsukuba Advanced Computing Center (TACC)

☆☆ Gaussian は量子化学計算プログラムのデファクトスタンダードであり、量子化学の専門/非専門を問わず化学者に広く使われる計算化学コードである。現在のバージョンは Gaussian98 である。

高性能なノード間スイッチが搭載されているがこのせっかくの並列機構が Gaussian ジョブが大多数を占める運用では十分に活用されていないと判断した。RS6000/SP は並列計算機システムとしての能力が高く、TACC として並列ソフトウェア開発用として利用されることを望んでいる。そこで、筆者等としてはこれらの Gaussian ユーザに対してアプリケーション利用を限定し、かつ計算速度としてより快適な環境の構築を目指し、計算要求をそちらに移行し RS6000/SP を高性能並列計算機として利用することを目指した QC GRID/Gaussian Portal の設計と構築を行った。設計の基本方針としては下記の通りである。

- スループットの重視
- コストパフォーマンスが良く、かつ安価なシステム構成
- グリッド技術を適用して、必要な計算サービスがいつでもどこからでも得られる
- 高性能スケジューラにより大規模ジョブに対しては RS6000/SP を選択する
- 課金のためアカウント管理、ログ等を採用する

本稿では Quantum Chemistry GRID/Gaussian Portal システムの設計と実装を述べる。既存の RS6000/SP システムとの性能比較についても報告する。また実際に Quantum Chemistry GRID/Gaussian Portal を産総研 TACC に導入し運用を開始した後の利用状況から得られた知見を報告する。次章からは Quantum Chemistry GRID/Gaussian Portal システムの基本概念、設計について述べ、3 章では実装例について述べ、4 章にて考察を述べ、最後にまとめを述べる。

2. Quantum Chemistry GRID/Gaussian Portal システム

2.1 基本概念

計算機ユーザーは自分にとって必要な計算の結果が、できる限り短時間に得られればよいのであって、どこで計算機を使うかが、どの計算機であろうか、OS が何であるかなどはこだわっていない。計算というサービス、計算の「結果」が得られれば満足する。我々はグリッドの考え方をを用いて、こうした計算サービスのあり方と将来像を描き、量子化学計算の応用技術とグリッド技術を融合させ、複数の量子化学計算資源、すなわち Gaussian、GAMESS、NWChem 等の量子化学計算アプリケーションやそれらのジョブ入出力データベース、文献データベース、ベンチマークデータベース、をグリッド技術によってユーザ認証、シングルサインオン、ファイルの共有等を実現しユーザーの利便性を考慮したシステム、QC GRID/Gaussian Portal システムを実験的に構築した。Gaussian Portal は QC GRID のフレームワーク上で実現した Gaussian

の Application Service Provider (ASP) として最初に実装されたものである。図 1 にシステム概略と処理の流れについて示す。

次節から QC GRID/Gaussian Portal を構成する個々の要素について詳細を述べる。

2.1.1 Gaussian Portal

Gaussian Portal は量子化学計算向け Portal の最初の実装例であり、

- Web インターフェース
- 結果 DB
- 知識 DB (入力 DB)
- 推論エンジン
- メタスケジューラ
- スケジューラを備えた複数の計算ノード

から構成されている (図 1)。

2.1.2 メタスケジューラ

メタスケジューラは以下の機能をもつプログラム群から構成されている。

- 入力ファイル評価機能
- 計算時間予測機能
- 検索時間予測機能
- 検索と計算の予測時間比較機能
- 高速検索機能
- 詳細検索機能
- 分子構造比較機能
- ジョブ制御機能
- 計算結果アーカイブ機能

その動作は、

- クライアント計算機から Web サーバーを経由して送られたジョブ要求の入力ファイルを読み込み解析し、計算手法を指定したコマンド行と分子構造から計算時間予測に必要な計算モデルと基底関数数を読み取り、結果 DB の検索に必要なデータ、高速検索用のコマンド行をキーワードとしたハッシュ生成、詳細検索用の個別オーバーレイのオプションからのハッシュ生成を行う。
- 読み取った計算モデルと基底関数数から計算時間を予測する。
- 予測計算時間と結果 DB を高速検索する時間を比較し、計算した方が速い場合にはジョブ要求を計算ノードのスケジューラに渡す。
- 結果 DB の高速検索時間が予測計算時間より十分小さいときは高速検索を実行する。高速検索では結果 DB に登録されている既存の結果の計算手法指定コマンド行からジョブ要求入力ファイルの計算手法指定コマンド行と一致するものを検索し、コマンドが一致した場合には分子構造比較機能により分子構造を比較し、一致した場合には結果 DB に登録されている既存の結果をユーザーに返す。
- 高速検索で一致する結果が無い場合、詳細検索実

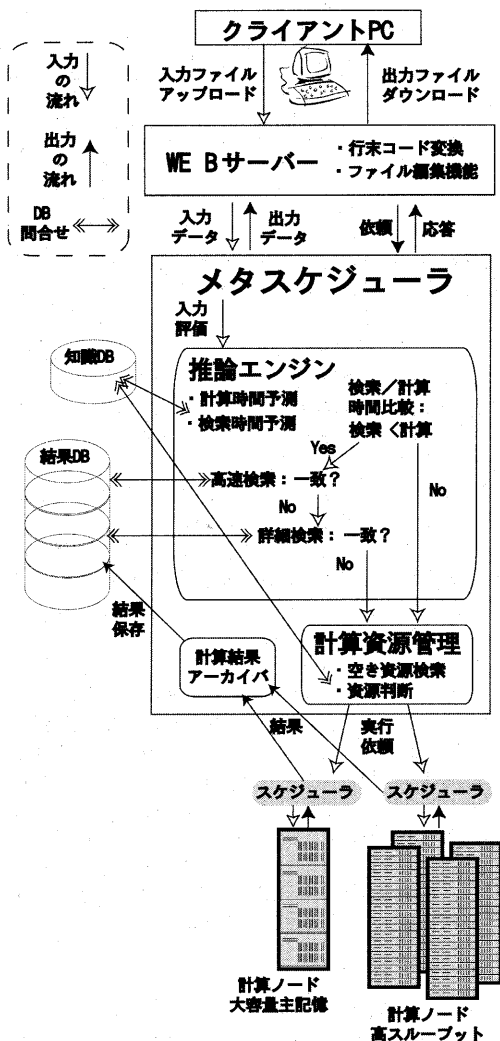


図 1 Quantum Chemistry GRID/Gaussian Portal システム

行時間と予測計算時間を比較し、計算した方が速い場合にはジョブ要求を計算ノードのスケジューラに渡す。

- そうでない場合は詳細検索を実行するかユーザーに確認し、類似性の高い既存の結果を検索するよう結果 DB に依頼する。詳細検索では、分子構造比較機能で類似の構造についても提示する。
- 検索実行後はいずれの検索でも類似性の高い結果を示し、ユーザーに結果 DB の結果を取り出すか、計算を実行するかを問う。
- 計算を開始した後はユーザーからのジョブの監視や中止などの制御を行う。実際のジョブは計算ノードのジョブ管理システム配下に管理されており、

メタスケジューラはユーザーに、様々なジョブ管理システムの差異を隠蔽し同一のインターフェースを提供する。

- 計算が完了した場合は、計算結果を Web インターフェースからユーザーに提供するとともに、結果 DB に登録する。その際に計算を依頼したユーザーに、結果を他者に提供する条件を問い、同様にその条件を結果 DB に保存する。結果が既存の計算結果から提供される場合には、出力ファイル中のユーザーアカウントなどを依頼者のものに置換し、結果 DB に検索がヒットし、依頼者に提供したとの情報を記録する。

2.1.3 入力ファイル評価機能

入力ファイルを読み込み、要求メモリ、スクラッチファイル、チェックポイントファイル、実行プロセス数、計算モデル、基底関数を評価し、ユーザーに提示する。アップロードした入力ファイルを解釈したポータル側のソフトウェアは下記のような項目を一覧としてユーザーに返す(表 1)。

ユーザーはこれらの表示を確認し、動作を指示する。また本機能では他に詳細検索で利用するために、組成式を作成し、一致する場合には分子の慣性主軸モーメントを算出し、分子主軸が一致するように回転、移動し、慣性モーメントや構造差の自乗和などから立体構造を比較する機能を持つ。

2.1.4 計算時間予測機能

計算モデルと基底関数から規格化された計算時間を求める関数が組み込まれており、さらに知識 DB に収められた計算資源の情報から、規格化された計算時間を実時間に近似し依頼ジョブの計算時間を予測する。

2.1.5 検索時間予測機能

知識 DB に収められた結果 DB の現在件数、検索 DB サーバの性能情報などから、ハッシュによる高速検索実行時間の予測や全文検索機能による詳細な類似性検索の検索時間を予測する。結果 DB は Gaussian Portal で計算された結果をすべて蓄積していくのでデータ件数が単調に増加して行くため検索時間が増大するが知識 DB にはデータ件数と検索時間、DB の配置と検索時間の関係が収められており、そのルールに

表 1 入力ファイルパーサーからの返り値
Table 1 return value of input file parser

| item | 内容 |
|-------------------------|-------------------|
| Method | モデル手法を表示 |
| Basis set | 基底関数系を表示 |
| Other keyword(s) | 上記以外の Keyword を表示 |
| Unidentified keyword(s) | 未解析の Keyword を表示 |
| Charge | 電荷を表示 |
| Spin mult | スピン多重度を表示 |
| # of atom | 原子数を表示 |
| # of function | 基底関数を表示 |
| Relative time for job | ジョブの規格化予想時間を表示 |
| Relative time for all | 全ジョブの規格化予想時間を表示 |

従って検索時間を予測する。

2.1.6 検索と計算の予測時間比較機能

ジョブ要求時点での結果 DB の予測検索時間と計算ノードから報告された現在の負荷情報も考慮した修正予測計算時間の大小を比較し、あらためて計算するか、検索を行うかを選択する機能を持つ。検索ステップでは完全一致のみの高速検索とユーザーに検索条件に適合しているかどうかを判断してもらう詳細検索の2段階の検索が可能である。

2.1.7 Web インターフェース

Gaussian Portal のユーザーはすべてを Web インターフェースを経由して作業を行なう。ユーザーは入力データを手元の計算機にファイルとして用意して Gaussian Portal の Web ページからアップロードし計算を依頼する。そのときクライアントマシンの OS 毎に異なる行末改行コードをユーザーが意識すること無く変換する機能を持つ。またジョブの監視や制御も Web ページから可能である。

利用モードとして「一般モード」と「エキスパートモード」の2種類を用意した。一般モードは、デフォルトの設定を利用し、ユーザはログイン以外の作業をほとんど要しない。ログインを行って、Gaussian の入力ファイルをアップロードし、ジョブを起動するだけであとは適宜計算ノードが選択され、実行を行う。システムでは、どのノードでジョブが実行されているか常にモニタしており、最も負荷の軽いノードにジョブが割り付けられる。このモードではユーザーは Gaussian Portal からの入力ファイル解釈と負荷予測結果の表示を確認した上で、実行結果を待つ。実行状態も Web ページから確認することができ、ジョブの終了を待って出力ファイルを確認する。出力ファイルは通常の標準出力と、エラー出力があり、正常終了している場合には標準出力ファイルから結果を手元の PC や Mac 等に回収する。一方、エキスパートモードは Gaussian に関して相当の知識と経験を有することを前提としジョブの実行に際しては実行スクリプトファイルを記述しなければならない。Gaussian の細かい制御文を記述することが可能である。また、Grid Port/Hot Page の技術を援用して、マルチサーバに対するファイルリストや他のファイル操作、編集作業が許されている。これらの操作により、きめ細かなジョブ制御、ファイル操作があたかも自分専用のシステムのように見せることが可能となる。この場合にはシステム側にインプットファイルを保存できるため、複数のジョブのチェイニング操作が可能となっている。

2.1.8 結果 DB

結果 DB は Gaussian Portal で計算された入力ファイルとその計算結果ファイルそのものをすべて蓄積していく。またデータを蓄積する際に検索用のインデックスを作成する。また結果 DB には、計算依頼者が結果の参照を他者に対して許可するか許可しないか、あ

表 2 Gaussian98 test ベンチマーク結果

Table 2 Benchmark results of Gaussian98 tests

| 年度 system | RS6000/SP CPU | GaussianPortal total time |
|-----------------|--------------------|------------------------------|
| RS 6000/SP | POWER3 250MHz | 488,263 sec. |
| Gaussian Portal | Pentium III 1.2GHz | 248,536 sec. |

る一定の期間後に参照を許すなどの参照ポリシーも登録される。さらに登録された結果が他者に参照される際には、元の計算依頼者の情報を隠す仕組みも実装されている。

2.1.9 知識 DB

計算モデルと基底関数から計算時間を予測する基準となる量子化学計算知識データを保持する。またグリッド上の計算資源のプロファイル保持し、どのような計算資源がどこに存在し、それらの負荷状況、時間変動、季節変動などを保持し、結果 DB の件数や検索に必要な時間を見積もるのに必要なデータなど計算資源に関する知識も保持する。メタスケジューラはこの知識を利用して規格化された計算時間を予測し、さらに知識データベースが保持する計算資源プロファイルから実時間を予測する。

2.1.10 計算資源

計算資源は目的・特性の異なる複数の計算ノード、高スループットを目的とした PC クラスタ、大容量メモリを搭載した SMP / NUMA システム、高速メモリ転送速度を持ちベクトル計算機などから構成されることを想定している。計算ノードはジョブ管理システムの管理下に置かれ、メタスケジューラからの計算依頼を受け取り、実際のジョブスケジューリングなどを行う。すなわちメタスケジューラにはパイプキューが置かれており、実際のキューは計算ノードに存在する。

3. 実装例

主たる計算資源として PentiumIII 1.2GHz 2way-SMP 主記憶 2GB ディスク 36GB の構成で 108 台を計算ノードとし、PentiumIII 1GHz 主記憶 1GB ディスク 300GB という 1 台のサーバノードから構成される PC クラスタを採用した。グリッド研究センターの技術提供により IBM と共同開発を行い TACC に設置し運用を開始した。

この計算資源の単体 CPU 性能を Gaussian98 Rev.A9 に付属のテスト入力データを実行し、既存の RS6000/SP と比較した結果を表 2 に示す。Gaussian Portal ではコンパイラの制限により 2GB を超える一時作業ファイルが作成できなかったり、一部の手法が実装されていなかったりするが、全体実行時間の比較への影響は無視できることを確認した。その結果、単体 CPU 性能は Gaussian Portal を構成する計算ノードが RS6000/SP の約 2 倍となった。

ソフトウェア構成はサーバーノードを除くノードは全て同じソフトウェア構成であり、OSはRedHat Linux⁴⁾、PGI コンパイラ実行環境⁵⁾、Gaussian、Open-PBS⁶⁾等がインストールされている。サーバーノードにはさらにApache⁷⁾、Globus¹¹⁾、GridPort、HotPage、OpenSSL⁹⁾、PostgreSQL¹⁰⁾がインストールされている。またサーバーノードは全ての計算ノードにNFSによってホーム領域300GBを提供し、定期的に保管するデータを格納する。各計算ノードでは作業領域20GBを確保しGaussian実行中に一時的に使用するデータを格納する。この領域は、各計算ノードのローカルファイルシステム領域に作成されている。%%実装についてもう少し詳しく書く、セキュリティ実装についてユーザー認証や安全な通信を実現するため、GlobusのGrid Security Infrastructure (GSI)を採用している。これによりシングルサインオンおよび委任 (delegation)、すなわちGlobus IDにて一度認証を受ければサーバー上にはユーザーのGlobus IDとローカルアカウントの対応表があり、ジョブはこの表のルールに従いGRIDを構成する全ての計算資源を適切なローカルアカウント権限で安全に利用することが可能となる。このためにはユーザーは自分の証明書を認証局 (Certificate Authority:CA) から発行してもらう必要があるが、今回の実装ではOpenSSLを利用し独自CA、self-signedのCAをTACC内に立ち上げた。現在の実装と運用体制ではユーザーの秘密鍵を作成するためのパスフレーズの第三者への漏えいを防ぐために、ユーザーはTACCまで実際に足を運んで端末の前まで来て自らの手でパスフレーズを打つ必要がある。ところがPortalを構築するのに利用したGridPort/HotPage実装では、ユーザーの秘密鍵をサーバー上に置くことになっており、セキュリティに不安が残る。パスフレーズの管理と秘密鍵の管理をもっと安全にユーザーが行え、かつ利便性も確保した実装が次の課題である。

計算機資源はGlobus Resource Allocation Manager (GRAM)を用いた。GRAMはNQS、LSF、PBSなど様々なサイト独自の資源管理ツールに対する共通のインターフェースを提供するため、計算資源の追加・拡張を容易にする。今回の実装ではローカルの資源管理ツールとしてOpenPBSを採用した。

4. 考 察

表3にGaussianの2、3月のCPU時間を示す。表3からQC Grid/Gaussian Portalの運用開始後、当初の意図とは異なりGaussianの計算要求がRS6000/SPからGaussian Portalに移行することは無く、CPU時間としてRS6000/SPに匹敵する時間がGaussian Portalにて純増となった。

このことはGaussian Portalを導入することで

表3 Gaussianの2、3月のCPU時間
Table 3 Gaussian CPU time in February and March

| 年度 | RS6000/SP | GaussianPortal |
|--------|-----------|----------------|
| 平成11年度 | 38,313 | — |
| 平成12年度 | 30,169 | — |
| 平成13年度 | 38,586 | 38,595 |

RS6000/SPでは実行しきれていなかった潜在的な計算需要を顕在化する事となった。また単体CPU性能ではGaussian PortalがRS6000/SPの約2倍となっているため、増加した計算需要は既存のRS6000/SPのCPU時間に換算すると2倍に相当する。RS6000/SPでGaussianのジョブの潜在要求を満たせていなかった要因としては、Gaussianが流せるシリアルキュー構成が全体の4割しか無かったこと、実行経過時間制限がS6000/SPでは200時間であったものがGaussian Portalでは無制限となったことが考えられる。Gaussian Portalは導入してから間もないため登録ユーザーはわずか7名に過ぎない。従って全体の稼働率は2、3月実績を平均して11%であった。それでもRS6000/SPの既存の計算需要を遥かに凌駕したジョブを処理した。このことから言えることは、これからのHPC資源を提供する計算センターの運用は、このことは常に言い続けられて来た事であるが、従来の細切れのキュー設定をしてはユーザーの需要にこたえることができず魅力を失っている。Gaussian Portalの用に計算需要の多い目的に対して専用の実行環境を用意し、ジョブ実行に関わる制限を無くすよう努力すべきである。

5. ま と め

グリッド技術を適用してユーザに簡便なインターフェースを提供し、誰でも簡単にGaussianを用いた計算することが可能となった。化学を専門としない者であってもPCやMac上のツールなどを使って基本データを作成し、それをポータルからジョブ投入することで計算結果を得ることが可能となった。ほとんど計算機に関する知識がなくても大規模な計算が可能となっている。これは、計算科学的手法を用いて、様々な物質、材料、創薬、生体反応等を解明するにあたって敷居を低くするが、計算の仮定および結果の妥当性に関しての検証が大問題として横たわっている。現在、計算時間の予測に利用している知識データベースを発展させることで、入力データが妥当なものの、そして計算結果が妥当なものを判断する際に有用なアドバイスをするシステムを実装する事を計画している。計算の前提と結果に対してのアドバイスシステムが完成した暁にはGaussian Portalは、計算による理論的検証を効率的に行なうためのツールとして、さらに有用性を高めるものと期待される。また計算資源の活用と

言う方向では、グリッド技術の適用を進め、様々な運用方針を持つ複数のサイトへの対応、計算コストを経済的に評価できるシステムを導入し、費用対効果を意識したスケジューリングを可能とすることなどにより柔軟かつ大規模な計算環境を提供することを目指す。今回は、産総研における試みを紹介したが、これらは容易に他所へ移植することが可能である。

謝辞 本研究は産業技術総合研究所先端情報計算センターの情報基盤研究開発室の業務の一環として実施された。先端情報計算センター長の寺倉博士には日頃よりご助力頂き感謝の意を表す。設計、デバッグ、実装に関してご議論いただいた建部修見博士ならびに田中良夫博士に感謝する。共同で作業を行った日本アイビーエム社の寒川光氏をはじめとするプロジェクトチームに感謝する。

参 考 文 献

- 1) Gaussian 98 (Revision A.9); M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle and J. A. Pople, Gaussian, Inc., Pittsburgh PA, 1998.
Gaussian Inc.: <http://www.gaussian.com/>
- 2) グリッド研究センター:
<http://unit.aist.go.jp/grid/>
- 3) 先端情報計算センター:
<http://unit.aist.go.jp/tacc/>
- 4) Red Hat, Inc. :
<http://www.redhat.com/>
- 5) The Portland Group, Inc. :
<http://www.pgroup.com/>
- 6) Veridian Systems :
<http://www.openpbs.org/>
- 7) The Apache Software Foundation :
<http://www.apache.org/>
- 8) M. Thomas, S. Mock, J. Boisseau, M. Dahan, K. Mueller, D. Sutton. The GridPort Toolkit Architecture for Building Grid Portals. Pro-

ceedings of the 10th IEEE Intl. Symp. on High Perf. Dist. Comp. Aug 2001

GridPort : <https://gridport.npaci.edu/>

9) The OpenSSL Project :

<http://www.openssl.org/>

10) PostgreSQL :

<http://www.ca.postgresql.org/index.html>

11) Globus: A Metacomputing Infrastructure Toolkit. I. Foster, C. Kesselman. Intl J. Super-computer Applications, 11(2):115-128, 1997.

GLOBUS project: <http://www.globus.org/>

12) HotPage :

<https://hotpage.npaci.edu/>