

Buffered Direct SCF 法の開発と性能評価

(産総研) ○稲富雄一^{*}、佐々木徹[†]、長嶋雲兵^{*}、村上和彰[‡]^{*} 産業技術総合研究所グリッド研究センター、[†] (株) アプリオリ・マイクロシステム[‡] 九州大学大学院システム情報科学

【序】Hartree-Fock(HF)法や密度汎関数(DFT)法に基づいた電子状態計算では、2電子積分の計算、操作に多くの計算機資源を費やしている。2電子積分の値はSCF計算の間に変化しないため、一度計算した積分値は再利用可能であるが、基底関数の数が1000を超える系になると、2電子積分ファイルをHDDなどの外部記憶装置へ保存することさえ困難になる。したがって、中規模もしくは大規模な系に対するHFあるいはDFT計算では、各SCFサイクルに必要な2電子積分計算を行うdirect SCF法を適用するほかない。ところで、膨大な量の2電子積分は独立に計算することが可能であるため、PCクラスタなどの並列計算機環境で高速に計算することが可能である。最近のPCクラスタでは、構成するPC1台あたりに1GB以上のメモリを搭載していることが多く、数1000軌道の系に対する計算では、未使用のメモリが多く存在する。そこで、空いているメモリ領域に2電子積分の一部を保存して再利用するbuffered direct SCF法を開発して、その性能を評価した。その結果、小さな保存領域(buffer size)を用いた場合でも、並列化効率が向上することが確認された。

【方法】2電子積分プログラムはvertical recurrence relation (VRR)とhorizontal recurrence relation (HRR)を組み合わせたhybridアルゴリズムを用いてC言語で記述した。また、計算速度向上と積分タイプ削減のために、contracted shell (CS)の並べ替えも行った。Buffered direct法では、最初のSCFサイクルでの積分計算で、与えられた大きさのbufferメモリに可能な限り2電子積分の値と4つの添え字を保存して、2回目以降のSCFサイクルで再利用するようにした。保存や再利用の際のオーバーヘッドをなくすために、閾値を超える積分を順番にbufferメモリに保存することとした。並列化は国際標準であるmessage-passing interface (MPI)を用いて行った。プロセッサが持つperformance counterを用いた積分プログラムの詳細な性能評価には、performance counter library (PCL)を使用した。

【結果と考察】積分を(pp,pp)から保存した場合と、(ss,ss)から保存した場合での台数効果の違いを図1に示す。これは、STO-3G基底関数を用いたグリシン15量体(Gly)₁₅の2回目のFock行列作成時間(2電子積分計算を含む)に対する台数効果で、buffer sizeは64MB/processorである。(ss,ss)から保存した場合と、(pp,pp)から保存した場合の結果を、それぞれ点線と破線で示してある。どちらの結果にもsuper-scalabilityが見られるが、(ss,ss)から保存した場合のほうが(pp,pp)から保存した場合よりも効率が高いことが分かる。これは、1つの積分値を求めるための計算コストが、(ss,ss)の方が(pp,pp)の場合に比べると大きいことによる。したがって、buffered direct法では、(ss,ss)から積分を保存することにした。

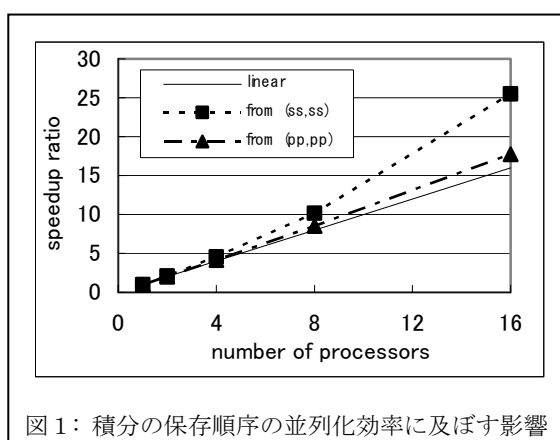
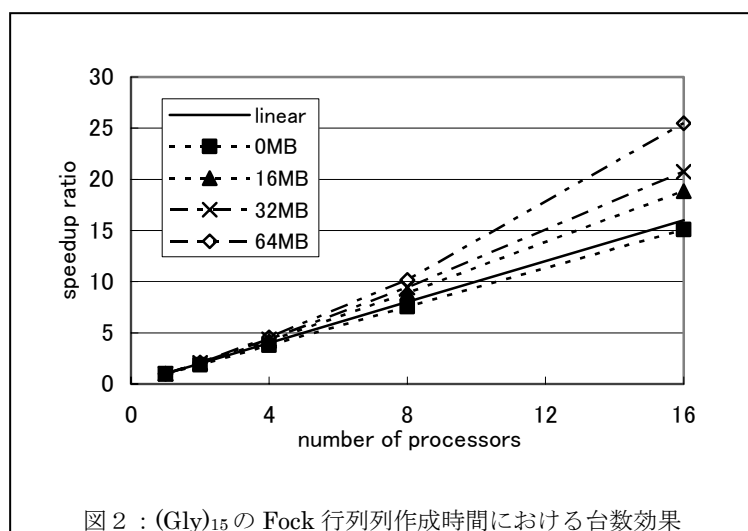


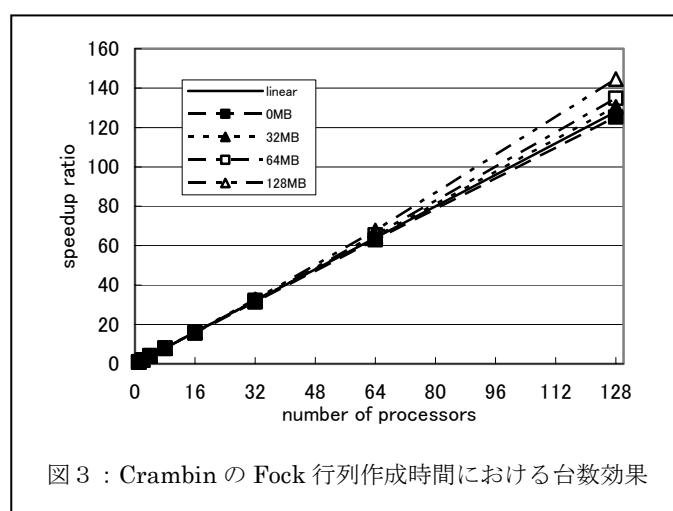
図1: 積分の保存順序の並列化効率に及ぼす影響



次に、小規模系（対象分子(Gly)₁₅、基底関数 STO-3G）に対する 2 回目の SCF サイクルにおける Fock 行列作成時間の台数効果を、いくつかの buffer size を用いて測定した結果を図 2 に示す。通常の direct SCF 法(buffer size=0MB)は非常に高い並列化効率を達成しているものの、理想的な場合(linear)よりも若干性能が落ちている。これは、並列化に伴い通信をする必要が生じることや、並列化できない部分が若干存在

するために起こる、一般的によく見られる現象である。一方で、buffered direct 法では super-linear な性能向上が見られた。Buffer size が 1 プロセッサあたり 16MB と小さい場合でも効果が表れているが、64MB の buffer size を用いた場合では 16CPU 使用時に 1CPU の場合の 2.5 倍を超える非常によい台数効果が得られた。

さらに、中規模系（対象分子 Crambin、46 アミノ酸残基、642 原子、基底関数 STO-3G）での結果を図 3 に示す。この系では、buffer size に 128MB/processor を用いた場合で、1processor あたりに全 non-zero 積分の 0.026% しか保存できない。しかし、台数効果に super-scalability が見られ、buffer size=128MB/processor で、128 CPU での計算を行った場合に、1CPU 使用時の約 144 倍の性能向上を示した。また、32MB/processor と非常に小さな buffer size を用いた場合でさえも、約 130 倍と linear scalability を超える速度向上をしていることが分かった。



【まとめ】剰余メモリ領域に 2 電子積分の一部を保存して再利用する buffered direct SCF 法の開発、および性能評価を行った。その結果、小さな buffer size を用いた場合でも、並列化効率を向上させる効果があることが分かった。現在開発中の 2 電子積分専用計算機システムでは、2 電子積分専用 LSI を並列に動作させて Fock 行列を高速に計算することを目論んでいるが、非常に多くの LSI を並列動作させる場合には、並列処理の際に表れる効率の低下が問題になる可能性がある。しかし、積分の一部を保存して再利用する buffered direct 法を用いることで、通常の並列化よりも高い並列化効率を達成できることが分かった。

本研究の一部は、科学技術振興費の総合研究「科学技術計算専用ロジック組み込み型プラットフォーム・アーキテクチャに関する研究」の支援による。また、すべての並列計算は産総研グリッド研究センター(GTRC)で行った。