

# DCB ライブラリを使用した効果的な負荷分散の検討

河合 直聡

東京大学情報基盤センター

## 1. はじめに

現在のスーパーコンピュータはマルチコアのノードを複数台高速なネットワークで接続した構成をとっている。このようなシステムを使用してアプリケーションを高速化する場合、ノード内とノード間の並列化を個別の実装で対応し (OpenMP+MPI)、ノード内およびノード間の負荷を均衡にする必要がある。多くのアプリケーションではこの多層な並列化構造に対応した場合に、全てのレイヤーで負荷を完全に均一にするのは困難である。例えば、主計算開始前の演算量の見積りが困難な場合 ( $\mathcal{H}$ 行列<sup>1</sup>) や、計算過程で空間的な演算量が変化する場合などが挙げられる。加えて、メニーコア CPU が広くスーパーコンピュータで利用されるようになり、ノード内の負荷分散がさらに困難となっている。このような演算量の不均衡は、スーパーコンピュータ使用時に期待通りの性能が得られないばかりか、待機コアの増加による無駄な電力消費にもつながる。

## 2. DCB ライブラリ

本研究では、プロセス毎の演算量に併せて割り付けるコア数を変更し、コア単位での演算量を均一化する Dynamic Core Binding (DCB) ライブラリ [1] を提案している。図 1(a) に示すように、一般的な OpenMP+MPI の並列化で、プロセス毎の演算量が不均一な場合は、プロセス毎に割り付けられるコア数が均等なため、コアレベルでの演算量も不均一となるが、DCB ライブラリを使用した場合はプロセス毎に割り付けるコア数を変更し、コアレベルでの演算量を均一化する。DCB ライブラリでは全てのコアを使用し、プロセス毎の演算量に併せて割り付ける“計算時間削減モード” (図 1(b)) と、演算量の多いプロセスに併せて演算量の少ないプロセスに割り付けるコア数を削減する“消費電力削減モード” (図 1(c)) を用意している。計算時間削減モードでは、演算量の多いプロセスに多くのコアが割り付けられるため、アプリケーション全体の計算時間短縮に繋がる。消費電力削減モードでは、ノード内の使用するコア数が減るため、割り付けられなかったコアはディープスリープ<sup>2</sup>状態に移行し、アプリケーション全体の計算時間をそのままに、消費電力削減が期待できる。

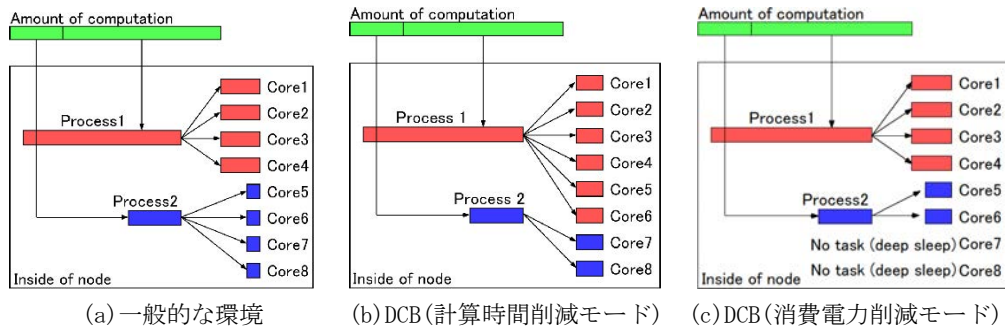


図 1 : DCB を使用した場合のプロセスに対するコア割り付けの一例

<sup>1</sup> Bebendorf Mario. “Hierarchical matrices”. Springer Berlin Heidelberg, 2008.

<sup>2</sup> <https://www.intel.co.jp/content/www/jp/ja/support/articles/000006619/processors/intel-core-processors.html>

### 3. 評価結果

図2にDCBの有無による効果を示す。図2はWisteria BDEC/01 Odysseyを最大24,336プロセス(ノード辺り4プロセス、6,084ノード)使用した結果を示している。評価ではLattice  $\mathcal{H}$ 行列[2]とベクトルの積にDCBライブラリを適用し、その行列ベクトル積を50回実施した。図1に示すように、DCB(計算時間短縮モード)の利用によって256プロセス(64ノード)使用時に36.2%、24,336プロセス使用時に8.5%の計算時間短縮を達成した。Oakbridge-CX(OBCX)に関しても同様の傾向を示しており、16プロセス(ノード辺り4プロセス、4ノード)使用時に40.6%、5,184プロセス(1,296ノード)使用時に4.5%の計算時間短縮を達成した。大規模並列環境でDCBの効果小さくなるのは、ノード間の負荷の不均衡によるものであり、この点に関しては今後の研究で解消予定である。消費電力削減モードでは、Odyssey 上では消費電力削減の効果が確認できず(消費電力が増加する結果となった)、OBCXでは5,184プロセスで46.7%の消費電力削減を達成した。Odysseyで消費電力削減の効果が得られなかったのは、デフォルトでretentionモード(CPUのディープスリープを有効にする機能)が無効になっているためであり、後の評価で49.1%(4,096プロセス、1,024ノード)の消費電力削減を達成している。

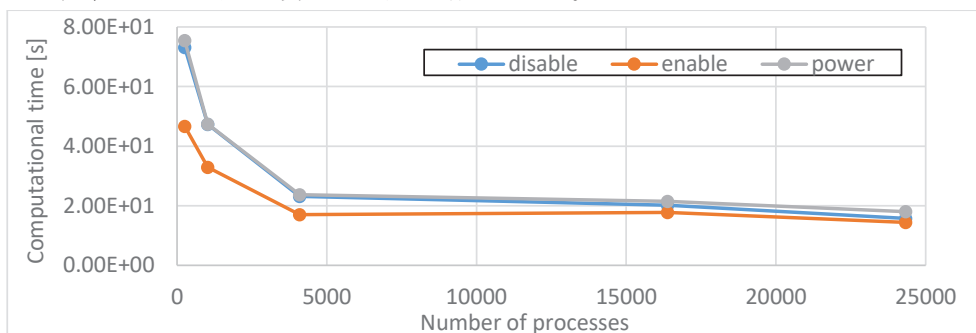


図2: DCBの評価結果(Wisteria-BDEC/01 Odyssey)

### 3. まとめ

本研究では、OpenMP+MPIで並列化された均一な負荷分散が困難なアプリケーションの動的負荷分散を実現するDynamic Core Binding(DCB)ライブラリの開発、評価を行っている。DCBライブラリはプロセス毎の演算量に併せてプロセスに割り当てるコア数を不均等にし、コアレベルでの演算量均一化を可能とする。大規模HPCチャレンジにて大規模並列環境で評価した結果、Wisteria-BDEC/01 Odysseyの6,084ノード使用時に8.5%、Oakbridge-CX(OBCX)の1,296ノード使用時に4.5%の計算時間短縮を達成し、Odysseyの1,024ノードで49.1%、OBCXの1,296ノードで46.7%の消費電力削減を達成した。

謝辞

今回、OBCXおよびWisteria-BDEC/01 Odyssey両システムの大規模HPCチャレンジを利用させていただいた。複数の大規模並列環境を利用させていただいたことで、様々な知見を得ることができ、より有効なライブラリ開発につなげることができた。ここに、本チャレンジ実施で様々な対応を頂いた皆様に感謝の意を表す。

### 参考文献

- [1] 河合,伊田,中島 "不均一なコア割付による動的負荷分散手法の検討" 応用数学会,2020年
- [2] A. Ida, "Lattice H-Matrices on Distributed-Memory Systems," 2018 IEEE International Parallel and Distributed Processing Symposium (IPDPS), 2018, pp. 389-398