

データベース処理のための 4AC-3 超並列プログラミング言語 MAPLLE における 動的負荷分散機構の評価

天野 浩文* 小野 剛† 堀淵 高照‡ 牧之内 顕文‡

*... 九州大学大型計算機センター

†... NTT九州移動通信網(株)

‡... 九州大学大学院システム情報科学研究科知能システム学専攻

1 まえがき

ハードウェア技術の進歩により、多数の PE (processing element) を有する超並列計算機が現実のものとなってきている。これをデータベース処理に生かすため、著者らは、コンパイル時には数とトポロジが不定であるようなオブジェクト集合に対して超並列処理を行うことのできるプログラミング言語の研究を行っている[2, 3]。

本稿では、そのような言語に必要となるオブジェクトの動的再配置機構[3]の性能について報告する。

2 オブジェクトの再配置戦略とその実装

これまでに提案した3つの再配置戦略の並列計算機 AP1000[4] 上のプロトタイプシステムにおける実装を概説する。

2.1 オブジェクト数の均等化

各 PE のオブジェクトの数を監視し、最大の値と最小の値の比が一定値を超えた時にオブジェクトの再配置を行う。再配置の開始はリダクション機能を用いて決定する。この場合、各 PE 上に存在するオブジェクト数のカウンタの値について最大値と最小値、平均値のリダクションを行う。最大値、最小値の比が一定の値を超えたとき再配置を開始する。

2.2 外部参照の削減

各 PE 上のオブジェクトを管理するテーブルの各エントリに、外部からの参照回数を記録するカウンタと、そのオブジェクトに最後にアクセスした PE の ID を記憶できるフィールドを付加する。管理下のオブジェクトに対し外部からのアクセスがあった時は、該当オブジェ

クトが登録されているエントリのカウンタの値をインクリメントし、その PE の ID を記録する。

あるオブジェクトに対する外部参照カウンタの値が一定の値を超えたら、当該オブジェクトの ID を再配置候補テーブルに登録しておく。負荷分散の開始は、再配置候補テーブルに登録されたオブジェクトの総数が一定の値を超えた時点とする。

2.3 特定の参照の局所化

参照が頻繁に発生すると管理者があらかじめ予測・指定した種類のリンクについては、そのリンクによる外部参照がなるべくおきないようにオブジェクトを再配置する。

今回の実験に用いたプロトタイプでは、局所化すべきリンクを発見したらその都度再配置候補テーブルに登録するようにした。

3 各再配置方式の性能

プロトタイプシステムに対して、負荷の不均衡を強制的に作り出してから再配置を行うことにより、各再配置方式をそれぞれ単独で適用した場合に性能がどのように変化するかを調査を開始した。これまでの結果はおおむね以下のように要約することができる。

3.1 オブジェクト数の均等化の効果

この手法は、ほぼ予想された通りの振る舞いを示した。オブジェクト数の最大値・最小値の比が 2, 3, 4, 5 倍となったときに再配置を行うと、その後の処理性能はほぼそれに見合う比率で向上する。再配置に要する時間は PE 数にほとんど無関係であった。

この手法は、単一のオブジェクトの再配置に必要な時間に比べて単一のオブジェクトに加えられる計算処理に必要な時間が無視できない程度に大きい場合に有効である。さらに、上記の実験結果が PE 数やオブジェクト数にほとんど依存しないことから、ごく少数の PE の上でこれらの時間を実測して比較することにより、かなり正確な性能予測が可能である。

Evaluation of Dynamic Load Balancing Mechanisms in MAPLLE: a Massively Parallel Programming Language for Database Processing

Hirofumi Amano*, Tsuyoshi Ono†, Takaaki Horibuchi‡, Akifumi Makinouchi‡

*... Computer Center, Kyushu University

†... NTT Kyushu Mobile Communications Network, Inc.

‡... Dept. Intelligent Systems, Kyushu University

3.2 外部参照の削減の効果

各 PE 上のオブジェクトの数が 1000 程度と小さい場合には、外部参照頻度を削減するために行ったはずの再配置が実際にはほとんど役立たなかった。これは、再配置先を決めるための情報が、最も頻繁にアクセスした PE ではなく最も最近にアクセスした PE という近似的な情報になっているため、および、再配置を行った結果、かえって外部参照が増えるようなケースを除外せずに一律に再配置を行っているためである。

一方、各 PE の持つオブジェクトの数を 10000 に増加させると、外部参照確率がかなり大きい (50% を越える) 場合に有効となるが、PE 数が小さい場合でも比較的よい値となっており、外部参照数の削減の効果も大きく現れるようになる。このような状況では、外部参照カウントと最後にアクセスした PE という近似的な情報である程度有効な再配置が可能であることになる。

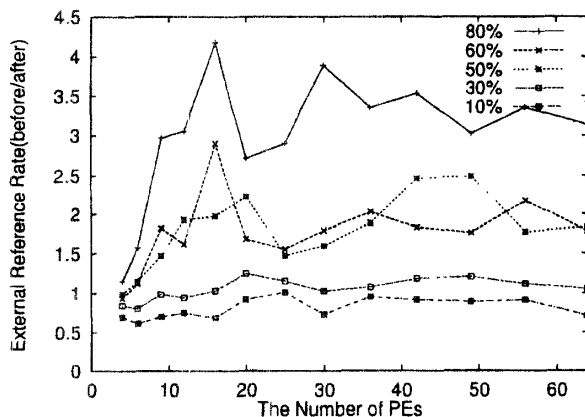


図 1: 外部参照の変化 (objects=10000, depth=7)

しかしながら、外部参照確率が 50% を越えるというのは実験のために強制的に作り出した極端な状況であり、実際にしばしば発生するとはあまり考えられない。むしろ、外部参照確率がある程度低い状況でもうまく動作する仕組みを考える必要がある。このために、現在、以下のような方法を検討している。

- 移動先の PE を決定するための情報をやや詳細化し、最後にアクセスした PE を複数回までさかのぼって記録しておき、ある程度参照元が一定しているオブジェクトのみ移動させる。
- 移動候補となったオブジェクトが参照している先 (比較的少数) を実際にたどり、その半分以上が局所参照となっている場合には、再配置を見合わせる。

3.3 参照の局所化に必要な情報収集コスト

再配置候補と異動先の決定に必要な情報の収集にかかるコストを比較するために、外部参照の削減と特定参照の局所化のそれぞれで必要とされる情報収集・維持のオーバーヘッドを含んだトラバース処理の実行時間を比較した。

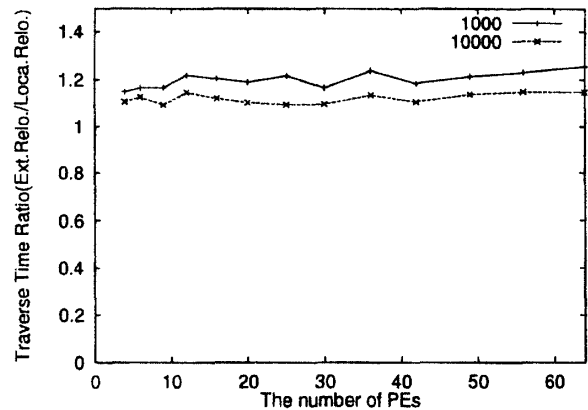


図 2: 外部参照の削減と特定の参照の局所化による性能向上の比較

図 2 に示すように、特定の参照の局所化のための情報収集のほうが、外部参照の局所化よりも 10~20% 程度コストが低いことがわかる。

したがって、「特定の参照」が管理者によって適切に設定されているなら、外部参照の削減のための再配置よりもより効率的に利用できる可能性が高い。

4 むすび

3つの手法のうち、外部参照の削減による再配置は予期したほどの性能向上が得られていないので、さらに改良を行う必要がある。

また、複数の手法を同時に用いて負荷情報を収集し再配置を行う場合には、相互に干渉を起こし悪影響を与える可能性が高い。そこで、今後はこれらの手法の相互干渉についても研究を進める予定である。

参考文献

- [1] Horie, T., Ishihata, H. et al.: "Design and Implementation of an Interconnection Network for the AP1000," *Information Processing '92*, Vol. I, pp. 555-561, 1992.
- [2] Imasaki, K., Ono, T. et al.: "Design and Evaluation of the Mechanism for Object Reference in a Parallel Object-Oriented Database System," *Proc. 1997 IEEE Int. Database Eng. and Applications Symp. (To Appear)*, August 1997.
- [3] Ono, T., Horibuchi, T. et al.: "Design of Dynamic Load Balancing Mechanisms for Massively Parallel Object-Oriented Databases," *Proc. Int. Symp. on Cooperative Database Systems for Advanced Applications*, Vol. 1, pp. 23-26, December 1996.