

Xcrypt 部門 発表

東北大学
工学部 情報知能システム総合学科
4年 奥田遼介

目的

- Perlベースのジョブ並列スクリプト言語
Xcryptを使用
- Linpackの実行パラメータ最適化
 - Nとノード数は固定

Linpack設定ファイル

HPLinpack benchmark input file

Innovative Computing Laboratory, University of Tennessee

HPL.out output file name (if any)

6 device out (6=stdout,7=stderr,file)

1 # of problems sizes (N)

32768 Ns

1 # of NBs

64 NBs

0 PMAP process mapping (0=Row-,1=Column-major)

1 # of process grids (P x Q)

8 Ps

8 Qs

16.0 threshold

1 # of panel fact

1 PFACTs (0=left, 1=Crout, 2=Right)

1 # of recursive stopping criterium

2 NBMINs (≥ 1)

1 # of panels in recursion

2 NDIVs

1 # of recursive panel fact.

2 RFACTs (0=left, 1=Crout, 2=Right)

1 # of broadcast

調整可能なパラメータ

□ 調整可能なパラメータは18個

- NB, PMAP process mapping, $P \cdot Q$, threshold, PFACT, NBMIN, NDIV, RFACT, BCAST, DEPTH, SWAP, swapping threshold, L1, U, Equilibration, memory alignment in double

□ どのパラメータが重要か？

- HPLの公式サイト¹⁾をよく読んでみる

□ 変更に意味のあるパラメータは少ない

- NB, $P \cdot Q$, NBMIN, BCAST, DEPTHを調整

1) <http://netlib.org/benchmark/hpl/tuning.html>

最適化の流れ

- 1CPU実行でのNB最適化
 - ノード固有の最適化
- P・Qの最適化
 - ノード分割パラメータの決定
- その他パラメータの最適化
 - NBMIN, BCAST, DEPTH

実行可能時間と計画

- 総実行時間
 - 1時間
- CPUコア数
 - 64以上かつ、MPI並列数の2倍以上
- 評価対象の問題サイズ
 - 対象環境で数十秒以上程度で終わるサイズ
- 実行回数(見積もり)
 - $1\text{分程度} \times 60\text{回} \times 2\text{並列} = 120\text{回程度}$

パラメータの実行見積もり

パラメータ	実行回数内訳	実行時間
NB	13種類 × 3回を64CPUでの並列実行	1分
P・Q	4種類 × 3回を2並列実行	6分
NBMIN	2種類 × 3回を2並列実行	3分
BCAST	2種類 × 3回を2並列実行	3分
DEPTH	2種類 × 3回を2並列実行	3分
単独実行	その時点の最適パラメータで3回	3分

総実行時間の60分で19分の実行を3回行い、
パラメータの収束を狙う

実行結果

問題サイズ

- N=32768, ノード数=64
- 理論性能 486GFLOPS

実行結果

0分 NB最適化開始

2分 P・Q最適化開始

NB=64決定

11分 P・Q最適化終了

228.4GFLOPS

P=2, Q=32

281.0GFLOPS

P=4, Q=16

285.7GFLOPS

P=8, Q=8

15分 NBMIN最適化終了

285.9GFLOPS

NBMIN=4

286.0GFLOPS

NBMIN=2

以後更新無し(4回目のNB最適化時に60分経過)

25%
改善

分析

- ❑ 理論性能の58.8%の理由
 - ❑ $286.0\text{GFLOPS} / 486\text{GFLOPS} = 58.8\%$
 - ❑ 富士通の実行効率は82.5%²⁾
- ❑ Nとノード数の大きさ
 - ❑ N^2 /ノード数が小さかった(4096)
 - ❑ 富士通実行時はメモリの限界となるNを決定
- ❑ BLASの性能
 - ❑ ATLAS3.9.25を使っていた
 - ❑ 富士通はFujitsuBlasを使用
- ❑ FlatMPIとHybridMPI
 - ❑ 通信のボトルネック
- ❑ バグ
 - ❑ NBが異常に小さい値に設定される

結論

- パラメータの選択が単純過ぎた
 - 40分の時間を無駄にしている
 - 複数を同時に変更する必要がある
- 理論性能に近づくには本体の変更が必要
 - 独自BLAS、HybridMPI

謝辞

- 指導教員を引き受けて下さった篠原歩教授と、機会を与えて下さった所属研究室の山口正洋教授に感謝致します。