ソフト工学用いた大粒度計算環境適応初期検討

An initial study of large-grained calculation adaptation using software engineering

山登庸次
Yoji Yamato　　yoji.yamato@ntt.com
NTT 株式会社　ネットワークサービスシステム研究所
Network Service Systems Labs., NTT, Inc.

1．はじめに

近年，多様ハードウェア利用が増えているが，活用の壁は高い．私達は，配置環境に合わせ，既存のコード変換等を自動で行い適切に動作させる，環境適応ソフトウェアを提案してきた[1]-[9]．これまで小粒度のループ文オフロード可否以外にも，大粒度機能ブロックの多様ハードウェアへの自動化で，抽象構文木を用いた機能ブロックオフロードを検証してきたが，類似度が非常に高い場合でないと利用できない問題があった．本稿では，大粒度機能ブロックの多様ハードウェアへの自動オフロード改善を対象とする．

2．機能ブロックのオフロード

本稿で取り扱う課題を示す．多様ハードウェアによるアプリケーション高速化は，高度な技術ノウハウを用いた手動製造が主流である．私達は，環境適応ソフトウェアを提案し，GPU 等へのオフロードを検証してきた．しかしこれまでは，遺伝的アルゴリズムを用いたループ文のオフロードや抽象構文木を用いた機能ブロックのオフロードの検証であり，ループ文オフロードでは，ハードウェア特性を生かしたライブラリ等での高速化には及ばなく，機能ブロックオフロードでは類似度が非常に高い場合でないと利用が難しかった．本稿は，フーリエ変換等の機能ブロックのオフロードの改善を対象にする．既存アプリケーションから，コード検知技術を用いて，オフロード可能なライブラリ等の実装がある機能ブロックか検知し，ある場合にそれを用いて処理する．

まず，コード分析し，オフロードできる計算処理機能ブロックかをコード検知技術で把握する．コード検知は，テキスト，トークン，抽象構文木，プログラム依存グラフ・意味，機械学習等の種類があり，DB 等検索対象と入力クエリ等の検索条件をマッチングし，類似コードがないかを検知する．ここで，機械学習のコード検知ツールを用いる．

コード検知に関連して事前に，オフロードできる計算処理の機能ブロックのコード，それをオフロード先で処理する場合の実装をコードパターン DB に保持しておく．この DB は，オフロード高速化に用いられるので，多様ハードウェアを提供するクラウド事業者等がインスタンスの利用活性化を狙い準備する事を想定している．機能ブロックの計算処理をハードウェアで処理する実装に関しては，別の方論文等で検討された物を用いる．

機械学習コード検知ツールで，オフロード可能な機能ブロックがあるか検索する．ここで，ない場合は，ループ文のオフロード高速化の試行に移行する．ある場合は，改めて抽象構文木コード検知ツールで機能ブロックの計算処理を行う実装を検索する．

抽象構文木の木構造を走査マッチングする．検索条件抽象構文木を検索対象抽象構文木上に部分木としてマッチするかどうかを判定する．抽象構文木部分木マッチングアルゴリズムで，検索条件抽象構文木を検索対象抽象構文木に対して走査し，部分木同型性を調べる．

機械学習系ツールと抽象構文木系ツールを組み合わせ，機械学習で低い類似度閾値等で多くの機能ブロックオフロード候補を発見できるようにし，抽象構文木でその機能ブロックの高速実装を発見して処理できるようにすることで，オフロード候補が少なくなるという問題を解決する．

GPU，FPGA，メニーコア CPU にオフロードできる計算処理の機能ブロックか判定できるため，発見された部分を GPU，FPGA，メニーコア CPU で処理する場合の実装に置換することで高速化する．種類として，GPU なら CUDA や OpenACC，FPGA なら OpenCL，メニーコア CPU なら OpenMP の実装である．実装は，MRI 系画像処理，離散 3 次元フーリエ変換等の計算処理機能ブロックで他論文等で手動改造で高速化が検討されてきた機能ブロックであり，専門家の今までのノウハウが詰まった実装と言える．

3．まとめ

本稿では，環境適応ソフトウェアの拡張として，ユーザがオフロード依頼するアプリケーションを，ループ文以外も分析して，フーリエ変換等の計算処理を行う機能ブロックを発見し，適切アルゴリズムで実装されたライブラリ等に自動オフロードする方式を提案した．今後は，複数アプリケーションでの評価を予定している．

参考文献

[1] Y. Yamato, "Proposal of Automatic GPU Offloading Method from Various Language Applications," ICIET 2021, pp.400-404, 2021.
[2] Y. Yamato, "Study and Evaluation of Automatic Offloading Method in Mixed Offloading Destination Environment," Cogent Eng., Taylor and Francis, 2022.
[3] Y. Yamato, "Proposal and evaluation of GPU offloading parts reconfiguration during applications operations for environment adaptation," J. Netw. Syst. Manag., Springer, 2023.
[4] Y. Yamato, "Study and evaluation for adopting environmental adaptation of low-resource devices," IEEE Access, 2024.
[5] Y. Yamato, "Study and Evaluation of FPGA Reconfiguration during Service Operation for Environment-Adaptive Software," Int. J. Parallel Emergent Distrib. Syst., Taylor and Francis, 2023.
[6] Y. Yamato, "Study and Evaluation of Optimum Location Deployment for Environment Adaptive Applications," Int. J. Parallel Emergent Distrib. Syst., Taylor and Francis, 2022.
[7] Y. Yamato, "Proposal and Evaluation of Adjusting Resource Amount for Automatically Offloaded Applications," Cogent Eng., Taylor and Francis, 2022.
[8] Y. Yamato, "Study and Evaluation of Automatic Offloading for Function Blocks of Applications," Automatika, Taylor and Francis, 2024.
[9] Y. Yamato, "Study and evaluation of optimal placement of applications in mixed environments including quantum computers for environmental adaptation," IEEE Trans. Cloud Comput., 2026.