

運用中の FPGA 再構成の検討

Study of FPGA logic reconfiguration during operation

山登庸次
Yoji Yamato

日本電信電話（株） ネットワークサービスシステム研究所
Network Service Systems Laboratories, NTT Corporation

1. はじめに

近年, IoT 等[1]-[3]新領域で, FPGA 等ヘテロハード利用が増えているが, 活用には壁が高い. 私は, 環境に合わせ, コード変換, 配置等を自動で行い高性能に動作させる, 環境適応ソフトウェアを提案し[4]-[7], 変換等取組んできた.

しかし, 現在までの検討は, サービス運用開始前にコードを分析して FPGA 等へ変換し, エッジ等に配置して, 価格や性能を改善して運用開始できるが, 運用開始後の利用特性変化等に応じて, 環境適応することはできない.

本稿では, 環境適応ソフトウェアの要素として運用開始前だけでなく, 開始後に, オフロードロジックを再構成することで変化に対応するための, 運用中再構成を検討する. 環境適応の再構成対象はオフロードロジック, 配置場所等多種あるが, 本稿は FPGA ロジックをターゲットとする.

2. 再構成必要性和 FPGA 着目理由

著者は, 以前に環境適応ソフトウェアのコンセプトを提案し, 全体像と要素技術を検証してきた.

全体像として, 運用開始前にまず, ユーザが指定したアプリケーションを, 分析し, コードの変換, リソース量の調整, 配置場所の調整, 検証を行ったのち, 運用を開始する. 次に, 運用開始後に, 利用特性等を分析して, 必要な再構成を行う. 再構成の対象は, 運用開始前と同様に, コード変換, リソース量の調整, 配置場所の調整である.

要素技術として, FPGA や GPU への変換等も検証してきた. GPU の際は, 遺伝的アルゴリズムにより, 殆どのループ文を対象に組合せを施行し, 1000 回規模測定を行って最適パターンを探索した. FPGA ではコンパイルに 6 時間以上かかるため, 性能測定回数を絞って探索する. ループ文 FPGA 自動オフロード時は, 算術強度やループ回数が高くリソース効率高いループ文に絞って, オフロードパターンを作り, 検証環境測定を通じて高速パターンの探索を行う.

現状を整理する. ヘテロジニアスハードに対するオフロードは手動での取組みが主流である. 私は環境適応ソフトウェアを提案し, ループ文等の FPGA や GPU 自動オフロード方式を検証しているが運用開始前の話で, 運用開始後の再構成については検討がされていない. そのため, 本稿では, FPGA オフロードロジックの, 利用特性に応じた再構成の方式を対象とする. 利用特性に応じて FPGA ロジックを再構成する例は, 商用クラウドでも例がなく, インパクトが大きいため, FPGA ロジックを対象とする.

3. 運用中 FPGA 再構成方式の検討

以前手法で, FPGA に自動オフロードできる. ただし, 運用開始前オフロードで用いる, 性能最適化用テストケース (複数のオフロードパターンで性能比較する際に性能比

較する項目) は, ユーザが指定した, 運用開始前の想定利用データを利用しており, 運用開始後に実利用されるデータから大きく離れる可能性がある. そのため, 運用開始後の利用が最初の想定と異なり, FPGA には別ロジックをオフロードした方が性能が向上する等の場合に, FPGA ロジックをユーザ影響低く再構成することを検討する. 再構成は, 異なるループ文オフロードに変える場合もあれば, 異なるアプリケーションのオフロードに変える場合もある.

FPGA 再構成は, 動的再構成, 静的再構成の 2 つがある. 断時間のユーザ影響度によって, FPGA 製造ベンダの提供する再構成手法を選択すればよいが, どちらでも断時間は発生することや, 別ロジックへの書換は動作確認の試験が必要なことから, 頻繁に再構成するべきとは考えておらず, 効果が閾値以上の場合だけ提案する等制限を設ける.

再構成は, 1 か月等一定期間のリクエスト傾向の分析から始まる. リクエスト傾向を分析し, 現在オフロードしているアプリケーションより処理負荷が高いか同等のものがあるかを把握する. 次に, 処理負荷が高いリクエストを, 想定利用データでなく実際に商用で利用されているデータを使って, FPGA オフロードの最適化試行を検証環境で行う. ここで, 検証により見つかった新しいオフロードパターンが現在のオフロードパターンより十分改善効果が高いかを閾値以上かそれ以下かで判定する. 閾値を上回る場合は, ユーザに再構成を提案する. ユーザ了承後, 商用環境を再構成するが, できるだけユーザ影響を抑え再構成する.

4. まとめ

運用開始後に, 利用特性に応じて, 適切な FPGA ロジックに運用中に再構成する, 運用中 FPGA 再構成方式を提案した. Intel FPGA PAC D5005 を用いて, 動作検証している.

参考文献

- [1] Y. Yamato, et al., "Security Camera Movie and ERP Data Matching System to Prevent Theft," IEEE CCNC 2017, pp.1021-1022, Jan. 2017.
- [2] Y. Yamato, et al., "Analyzing Machine Noise for Real Time Maintenance," ICGIP 2016, Oct. 2016.
- [3] Y. Yamato, "Experiments of posture estimation on vehicles using wearable acceleration sensors," IEEE BigDataSecurity 2017, pp.14-17, May 2017.
- [4] Y. Yamato, "Study of parallel processing area extraction and data transfer number reduction for automatic GPU offloading of IoT applications," J. Intell. Inf. Syst., Springer, 2019.
- [5] Y. Yamato, "Study and Evaluation of Automatic GPU Offloading Method from Various Language Applications," Int. J. Parallel Emergent Distrib. Syst., Taylor and Francis, Sep. 2021.
- [6] Y. Yamato, "Study and Evaluation of Improved Automatic GPU Offloading Method," Int. J. Parallel Emergent Distrib. Syst., Taylor and Francis, June 2021.
- [7] Y. Yamato, "Proposal of Optimum Application Deployment Technology for Heterogeneous IaaS Cloud," WCSE 2016, 2016.