

平成 30 年 5 月 25 日現在

機関番号：12613

研究種目：基盤研究(C) (一般)

研究期間：2014～2017

課題番号：26330143

研究課題名(和文) 計算アクセラレータによる高効率なネットワークサービス基盤の実現

研究課題名(英文) Implementing Efficient Infrastructure for Network Services by Utilizing Computational Accelerators

研究代表者

矢崎 俊志 (Yazaki, Syunji)

一橋大学・情報基盤センター・助教

研究者番号：60454138

交付決定額(研究期間全体)：(直接経費) 3,700,000円

研究成果の概要(和文)：効率の良いクラウドサービスなどへの応用を目指して、多数の計算コアや高速メモリを持つ計算アクセラレータを用いた高効率なネットワークサービス基盤と評価用ウェブアプリケーション FEHA (Front-end Environment for Hands-on Activities) の開発を行った。また、開発した基盤とFEHAを大学における多人数講義に適用し、評価用ワークロードシナリオ作成環境を構築した。実装した基盤とFEHAを約50から150名の参加する大学の講義で使用し、要求されたリクエストが相対的な高負荷時においても遅延なく処理できていることを示した。

研究成果の概要(英文)：In this research, I proposed an efficient infrastructure model for network services that is a base of the cloud computing. I designed the mode to utilize computational accelerators that have many integrated processing units and high bandwidth memory. I also developed a web application that named Front-end Environment for Hands-on Activities (FEHA) to evaluate the proposed model and, implemented an environment to collect request histories. The history can be used to make work-load scenarios for assessing infrastructures that provide practical web applications in actual services. We deployed FEHA on an implemented infrastructure that refers to the proposed model and used it in lectures at a university. About 50 to 150 students joined to the lectures. As a result, I showed that my infrastructure and FEHA processed all requests without significant latency even under a high-loading situation.

研究分野：高性能計算

キーワード：ネットワークサービス基盤 計算アクセラレータ ウェブアプリケーション

1. 研究開始当初の背景

近年、クラウドサービスの利用が進み、情報サービスの利用形態が大きく変わりつつある。従来、端末にインストールされたアプリケーションにより提供されてきた機能が、ウェブサービスとして提供される事例 (SaaS, Software as a Service) が多くなってきた。ウェブメールなどがその一例である。このようなクラウドサービスの利用を前提とした社会サービスの需要も増え、組織間の業務提携・データ連携などの活性化も相まって、ウェブシステムをはじめとする情報ネットワークサービス全体に対する大量アクセスや大容量データ転送の需要は日々高まっている。

この需要増に対応した高速・大容量ネットワークサービス基盤を実現するために必要な機器のコストも日々増大している。一般向け情報サービス基盤においては、家庭用の10倍～100倍 (10 Gbps ~ 100Gbps) の要求に耐えうるバックボーンネットワークが主流となっている。特定の組織内においても、10 Gbps 以上のネットワークが用いられるようになった。

主に大規模な通信・クラウド事業拠点となる大規模データセンタのネットワークサービス基盤においては、高コストで高性能なデバイスである ASIC (Application Specific Integrated Circuit) , FPGA (Field Programmable Gate Array) , CAM (Contents Addressable Memory) などを利用した専用機器を用いるのが一般的である。

一方で、利用者に対してより高いサービス品質を保証するために重要なエッジコンピューティングにおいては、大規模データセンタとサービス利用者の間に中・小規模データセンタなどのサービス提供拠点を配置する必要がある。この様な中・小規模サービス拠点においては、前述の高価な専用機器だけでなく一般的な PC サーバなどを活用することで、合理的なコスト/性能比を実現する必要がある。

PC サーバによるネットワークサービス基盤の構築においてネットワーク実装の階層化モデルにおける第3層である IP (Internet Protocol) アドレス層以下では、マルチコアプロセッサを活用した並列化によるルーティング処理などの高速化は既に標準である。近年、よりコスト性能比の良い方法として、GPU (Graphic Processing Unit) が持つ、高速メモリや数千個の実数演算回路を計算アクセラレータとして活用することで、IP アドレス検索、暗号化処理、あいまい比較などの計算バウンドな処理を高速化している例もある [1][2]。第4層以上 (トランスポート層) においては、SSL (Secure Socket Layer) などでも多用する暗号処理なども同じく計算バウ

ンドであるため、これらを GPU で高速化している例もある [3]。一方で、最上位層 (アプリケーション層) に位置するウェブサーバ等の、アプリケーションサーバを実現するサービスプログラムに対しては、20 CPU コア程度のマルチコアプロセッサを利用して、リクエストに対するプロセスまたはスレッド並列化が行われているが、下位層における高速化度合いと比較して不十分である。

ネットワークサービス基盤全体として、よりコスト性能比の良いシステムを実現するためには、下位層の高速化に見合った効率的なリクエストの並列処理方法を、上位層においても確立する必要がある。

<参考文献>

- [1] Sangjin Han, Keon Jang, KyoungSoo Park and Sue Moon, "PacketShader: a GPU-accelerated software router," ACM SIGCOMM Computer Communication Review, Vol.40, No.4, pp.195-206, 2010.
- [2] P. Sundaresan and S. Venkateshprasad, Y. Muralitharan, R. Parthasarathi, "gR: A GPU-based Router", Available at <https://protected.fatcow.com/hipcor/hipc2012/documents/SRSPapers/Paper49.pdf>, 2012.
- [3] Johannes Gilger, Johannes Barnickel, and Ulrike Meyer, "GPU-Acceleration of Block Ciphers in the OpenSSL Cryptographic Library", Information Security, Vol.7483, pp.338-353, 2012.

2. 研究の目的

本研究計画では、アクセラレータとして GPU および MIC (Many Integrated Cores) を実装した Co-Processor ボードを用いて、それぞれが得意とする処理の特徴を考慮して、ネットワーク実装モデルにおける第3層以上を実装するサービスプログラムの高速化を行う。これにより、コスト性能比の良いネットワークサービス基盤を実現することを目的とする。

3. 研究の方法

ネットワークサービス基盤に対するリクエストは、計算バウンドなものと、I/O バウンドなものに大別できる。前者は、リクエストあたりの計算量が多く、計算機の演算能力を主に消費する。後者は、計算機のディスクやネットワーク装置の処理能力を主に消費する。

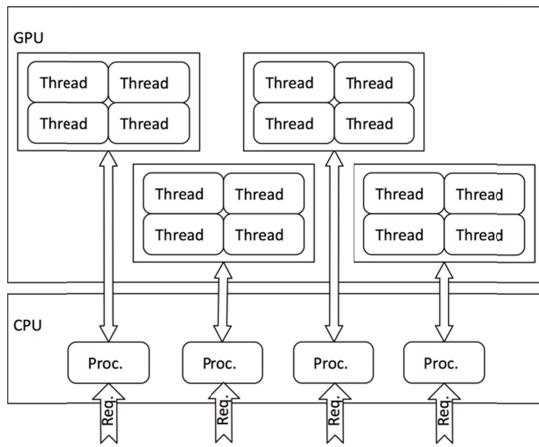


図 1 : GPU による分散処理モデル

本計画では、これらをそれぞれ異なるモデルで扱う。計算バウンドなリクエストに対しては、GPU を活用し、図 1 に示すように、CPU (Central Processing Unit) 上の各プロセスから起動された多数の GPU スレッドを用いて、処理を GPU にオフロードする仕組みを実装する。これは、既存研究ですでに提案されている。計算バウンドな処理の具体例として、ウェブサービスの処理に多くに含まれる文字列のマッチングや検索などの処理がある。

一方、I/O バウンドな処理については、MIC (Many Integrated Core) と HBM (High Bandwidth Memory) を実装した Co-Processor ボードを用いて効率化を行う。本研究では、主にこのモデルを研究対象とする。I/O バウンドな処理の例としては、ウェブサービスにおけるリクエスト処理が挙げられる。近年のネットワークサービス基盤は、ユーザやプログラムのインタフェースとして HTTP (Hyper-Text Transfer Protocol) を採用しているものが多く、効率化によって大きな改善が見込める。

近年の Co-Processor ボードは、一般的な CPU 単体よりはるかに多い、60 個の CPU コアと、同じく一般的なメモリより高速な HBM を搭載しているものもある。このような Co-Processor ボードは、個々の CPU コアの処理能力は既存の CPU コアのものよりも低いですが、ウェブサービスのように、小さなリクエストを大量に処理する場合には、HBM と合わせたディスク I/O 最適化と合わせて、高い並列化効率が期待できる。

具体的な仕組みとしては、図 2 に示すように、CPU 上のプロセスを Proxy (代理) として動作させ、Proxy が受け取ったリクエストを Co-Processor ボード上の小さな CPU コアに分散処理させる。これにより、より多数のプロセッサと高速なメモリを活用した大量プロセス並列処理基盤を実装する。

加えて、本計画では実装したネットワークサービス基盤を評価するための環境を構築す

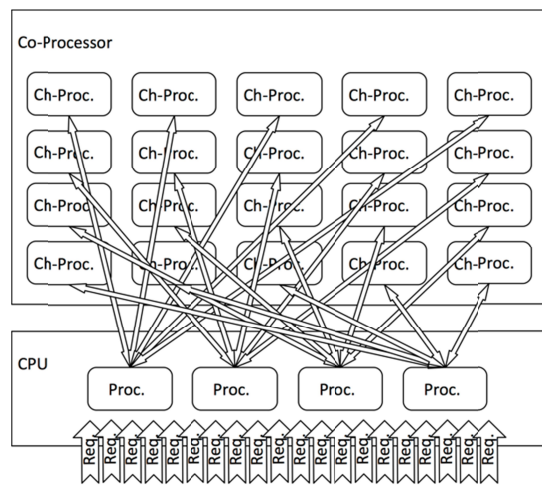


図 2 : MIC 型 Co-Processor による分散処理モデル

る。ネットワークサービスに対するリクエストの頻度や処理量は、サービス内容や運用状況によって大きく異なる。主に、開発の過程においては、ネットワークサービス基盤を構成する個々の要素、例えば、特定の演算に要する時間や特定機器へのリクエスト処理時間などを評価するためにマイクロベンチマーキングが行われる。これらは既存ツールで実施できる。

評価をより実運用に近い環境で行うためには、ネットワークサービス基盤へのリクエストパターンをシミュレーションするツールの開発や、実運用データに基づく評価シナリオの作成が必要である。本計画では、大学での多人数講義に活用できるウェブアプリケーションを開発し、その稼働データを収集し、評価シナリオの作成を行う。

4. 研究成果

研究実施期間中における本研究の主な成果は以下のとおりである。

1. 評価用ウェブアプリケーションの実装と評価用ワークロードシナリオ作成のためのデータ収集
2. 一般 PC を用いたプロセス並列化ウェブサービス基盤の実装
3. Co-Processor ボードを用いたプロセス並列化ウェブサービス基盤の実装

1. について、評価用ウェブ・アプリケーションとして、FEHA (Front-end Environment for Hands-on Activities) を開発した。FEHA は、既存のスーパーコンピュータなどを計算のバックエンドとして改変なく利用し、同時に、利用者に簡易的なプログラミング環境を提供するウェブアプリケーションである。

近年のウェブアプリケーションは、多数のリクエストが同時並行で発生する状況においても、サービスの応答性を維持するため、単

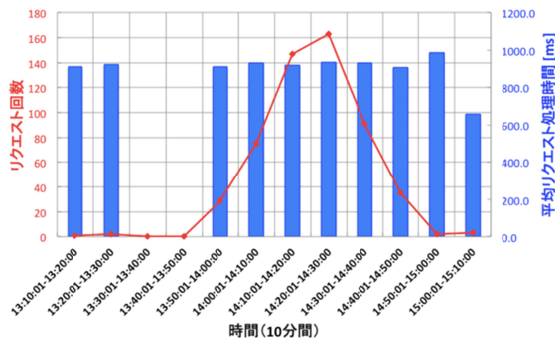


図 3：一般 PC を用いたウェブサービス基盤上で動作させた FEHA で収集したリクエスト履歴とその応答時間．リクエスト回数（左軸）の増減に関わらず平均リクエスト処理時間（右軸）に大きな変化がない．要求されたリクエストがどの場合においても遅延なく処理できている事を示している．

ープロセスによるイベントループ処理と非同期ファイル I/O 処理を組み合わせることで、リクエストに対する処理をノンブロッキングで行う．FEHA の実装においても、現実のウェブアプリケーションと同様の処理方式を採用した．

実装した FEHA を、約 50 名から 150 名が参加する大学の講義に複数回適用し、ネットワークサービスへのリクエストが短時間に集中的かつ非機械的に発生する環境を再現し、リクエスト履歴を収集した．このリクエスト履歴をもとに、現実のウェブアプリケーションを想定したワークロードシナリオが作成可能になった．

FEHA については、そのプログラミング環境提供方法としての実用性が認められ、査読付きのジャーナル論文および国際会議の発表論文として採録された．

2. については、一般 PC (Personal Computer) 程度の性能を持つ計算機上に、プロセス並列方式でウェブサービス基盤を構築した．本研究では、プロセス並列方式として図 2 に示したモデルを実装した．このモデルでは、上述のように、サービスを代表するプロセスを Proxy として動作させ、実際の処理を他のプロセスに移譲することで並列処理を実現する．また、実装した基盤上で動作させた FEHA を用いて、約 50 名が参加するプログラミング講義中に発生するリクエストを、どの程度の応答時間で処理できるかを測定した．

図 3 に実験結果を示す．図から、リクエスト回数の増減に関わらず平均リクエスト処理時間に大きな変化がないことが読み取れる．これは、本研究で構築した基盤において、要求されたリクエストが相対的な高負荷時においても遅延なく処理できていることを示している．

3. については、2. の実験結果を踏まえ、高性能計算向けの Co-Processor ボードを使用して、より並列度の高いウェブサービス基盤を構築した．実験用に選定した Co-Processor ボードは、前述のように、MIC と HBM の両方を実装したものである．これらを活用することで、MIC による小さな大量のリクエスト処理を、HBM を用いた高速なファイルキャッシュと組み合わせることで、効果的なリクエスト処理が期待できる．

研究期間において、3. についてはハードウェア構成まで完了したが、2. で構築したウェブサービス基盤を 3. のハードウェア構成で効率よく動作させるためのソフトウェアの改良が完了していない．これは今後の課題である．

5. 主な発表論文等

〔雑誌論文〕(計 2 件)

1. [Syunji Yazaki](#), Hideaki Tsuchiya, and Hiroaki Ishihata, "Development of A Web-based Front-end Environment to Aid Programming Lectures on Unix-like Systems," IPSJ Journal of Information Processing, 査読有, Vol.59, No.5, May2018. (in printing)

〔学会発表〕(計 2 件)

1. [Syunji Yazaki](#), Takeshi Kikuchi, Hideaki Tsuchiya and Hiroaki Ishihata, "FEHA: an Adaptive Web-Based Front-End Environment to Support Hands-On Training in Parallel Programming," Proceedings of the Future of Education 2016, 査読有 pp.166-170, 2016.
2. 矢崎 俊志, 石畑 宏明, "ハンズオン講義・デモンストレーションを支援する計算機システム向けフロントエンド環境の提案-並列処理の講義への適用-", 情報処理研究会報告 第 148 回 HPC 研究会, 査読無, Vol. 2015-HPC-148, No. 25, pp. 1-6, 2015.

6. 研究組織

(1) 研究代表者

矢崎 俊志 (YAZAKI, Syunji)
一橋大学・情報基盤センター・助教
研究者番号：60454138