

TSUBAME 共同利用 平成 29 年度 学術利用 成果報告書

利用課題名 HPC を利用した自然言語処理技術の研究

英文: High Performance Computing for Natural Language Processing Technology Research

利用課題責任者

鳥澤 健太郎

所属

国立研究開発法人情報通信研究機構 ユニバーサルコミュニケーション研究所

データ駆動知能システム研究センター

<http://www2.nict.go.jp/direct/>

邦文抄録(300 字程度) 質問応答や対話等の高度な自然言語処理技術には、大量のテキストの意味的に深い分析が必要であるが、数十億規模の Web 文書の分析には、多数の計算機を用いた並列実行が必須である。本課題では、各種の分析プログラム群を TSUBAME 上で多数並列実行し、大量の Web テキストを高速に処理することを目指した。その結果、大規模 Web 情報分析システム WISDOM X で必要となる分析を、1 日 1 億 Web ページ以上の速度で実行できた。また、200 億レコード(1.5TB)超のデータに対して、対話システム WEKDA で用いられる深層学習プログラムを適用し、最大 100 以上の GPGPU を使用し、16 時間で処理できた。さらに、ニューラルネットワークの自動分割による、複数 GPU を用いた並列処理を実現し、同じく WEKDA で用いられる深層学習プログラムの学習を約 2 倍高速化した。

英文抄録(100 words 程度) Advanced natural language processing applications including question-answering and dialogue systems require semantic deep analyses on a large amount of texts. Such analyses should be distributed on many compute nodes because the semantic analyses consist of computationally heavy programs. In this project, we introduce our middleware RaSC and tuned the configurations for TSUBAME. As the result, we processed 100 million web pages per day with analysis programs for WISDOM X. We also processed 20 billion records in 16 hours with our deep learning programs for WEKDA using more than 100 GPUs. We also automated automatic model parallelism on multiple GPUs and doubled the training speed of the deep learning programs.

Keywords: 自然言語処理, 大規模情報分析, テキスト分析, モデルパラレル

背景と目的

質問応答や対話等の高度な自然言語処理技術には、大量のテキストの意味的に深い分析が必須である。例えば、情報通信研究機構が開発している大規模 Web 情報分析システム WISDOM X¹ や対話システム WEKDA² には、40 億以上の Web ページの解析結果が用いられている。こうしたアプリケーションのためのテキストの分析は、計算機負荷が高く、数十億規模の Web 文書の分析には、多数の計算機を用いて並列に分析を実行することが必須である。

そこで本課題では、当機構の持つ分析プログラム群を TSUBAME 上で多数並列実行し、大量の Web テキストを高速に処理することを目指した。

概要

当機構に蓄積された Web テキストについて、TSUBAME 環境で各種の分析プログラムを適用した。本課題では、以下に示す 3 つのケースを実施した。
ケース 1) ストリーミング処理によるプログラム連携
当機構で用いている分析プログラムは、速度や使用メモリ量が様々であり、複数の分析プログラムを連携させるにあたって、単純に全てのプログラムを同一プロセス数だけ起動してパイプラインを構成すると、メモリ消費量等の点で非効率である。従って、当機構が開発したミドルウェア RaSC を用い、各分析プログラムに最適な数のプロセスを起動してプールし、それらをストリーミング通信によって連携する構成とした。

バッチジョブの中でこの構成を実現するため、バッチジョブの冒頭で、分析プログラムの実行環境構築と、

¹ <http://wisdom.nict.jp>

² <https://www.nict.go.jp/press/2017/10/24-1.html>

それらのプログラムのプロセスをプールする RaSC サービスの起動等を行った。RaSC を用いる場合、全体の実行を管理する中央サーバによって、分析プログラムの複数の実行ノードへの配備や、プログラムの死活監視等を一括して行える。しかし、TSUBAME 環境においては動的に計算ノードが割り当てられるため、複数のジョブにまたがって持続的に中央サーバを起動しておくことが難しい。そのため、各ジョブが一つの計算ノードを用いるものとし、その計算ノードに中央サーバと全ての分析プログラムを設置・起動してから、分析対象データを送信することとした。

ケース 2) 深層学習プログラムのバッチ実行

本課題では、学習済みの深層学習モデルを用いて、Web テキストに対する推論を行うプログラムを実行した。深層学習のためのライブラリには、実行時に GPGPU のための実行コードを生成し、コンパイルを行うものがある。コンパイルされたコードは、ディスク上の一時領域に保存される。本課題で用いた深層学習プログラムでは、深層学習のためのライブラリに Theano 及び PyTorch を用いたが、前者はユーザのホーム領域を一時領域として、GPGPU 用コードの出力とコンパイルを行う。多数の計算ノードが同一のパスを参照することになるため、ロック開放待ちやサイズの小さいファイルが多数読み書きされるため、パフォーマンスが低下する。そのため、コンパイル結果の配置は、ノードごとに独立したキャッシュ領域に書き込むように設定して実行した。

ケース 3) モデルパラレルによる並列化

上記で用いたニューラルネットワークが、並列実行可能な分岐構造を持つことから、ニューラルネットワークを分割し、各部を異なる GPGPU 上で計算するモデルパラレルの自動化を試みた。このため、ニューラルネットワークを構成する各計算モジュールの計算時間をプロファイリングするとともに、最適な分割方法をスケジューリングによって求めるようにした。

結果および考察

上述のケース 1) においては、対象となる全データは圧縮状態で数百 TB 規模となるため、本課題ではその一部を対象とした上、200 万 Web ページ相当のデータ (約 30GB) を 1 単位として TSUBAME にコピーし、1

ジョブとして実行した。各ジョブ内においては、RaSC を用いた並列実行が行われるため、資源リソースには多数のコアを持つ `f_node` を用いた。最大でおよそ 50 ジョブが同時実行され、一日あたり 1 億 Web ページを超える分析が実現された。

またケース 2) においては、対象となる Web テキスト約 1.5TB (200 億レコード) を全て TSUBAME のストレージにコピーした上で、学習済み深層学習モデルによる推論を行った。また、全データを 200 に分割し、それぞれを 1 つのジョブとして実行した。各ジョブは GPU を 1 つのみ使用するため、資源タイプには `q_node` を用いた。その結果、最大で 100 ジョブ以上を同時稼働することができ、およそ 20 時間で全ジョブを完了した。

ケース 3) については、WEKDA で用いる深層学習プログラムの学習を、2 倍以上高速化することに成功した。同一の計算機上の GPGPU を用いるため、資源タイプには 4 枚の GPGPU が利用可能な `f_node` を用いた。高速化効果には、GPGPU 間の通信速度が極めて大きく影響するため、TSUBAME において GPGPU が NVLink で接続されていることが大きく寄与したと思われる。

まとめ、今後の課題

TSUBAME の計算機リソースが、ストリーミング処理による分析プログラム連携を行うケースや、深層学習プログラムを用いる行うケースなど、幅広い状況で大量データの処理に有用であることが確認された。

一方で、TSUBAME 上に配置不可能な大規模データを処理するため、データを分割してそれぞれをバッチジョブで処理していることから、入出力データのディスク読み書きのコストを生じている。今後、TSUBAME 環境に起動した分析プログラムに対し、ストレージを介することなく、直接ネットワークでの入出力データの送受信を行うよう拡張することを検討している。

また本課題で実施したモデルパラレルの自動化は、1 台の計算機上の GPGPU を用いることを前提としたものであり、スケーラビリティに限界があった。今後、TSUBAME が多数の GPGPU 計算機を高速に結合していることを生かし、複数の GPGPU を用いるように拡張していく予定である。