機械学習向けのコンピュータシステムの構築に向けた AI ワークロードの特徴分析

高山沙也加　（指導教員：小口正人）

1　はじめに

AI を用いたアプリケーション利用の増加に伴って，CPU と比べて処理性能が高く電力消費の激しい GPU の利用が進み ICT システムの全体電力が増加傾向にある．特に近年では，プロセッサの高負荷による大量の電力消費がもたらす環境面への影響が懸念されている [1]．そのため，システム性能を落さない範囲での ICT 電力の削減が望まれ，システム電力削減とアプリケーション性能の向上のバランスを取り，システムを効率良く稼働させることがますます重要になる．ワークロードという計算機に対する負荷または負荷アプリケーションを使った実行時性能やリソース情報から効率的にハードウェアを設計･制御する手法は既に用いられている [2]．この制御にあたって CPU 時間やデータベースの変更数，応答時間などの情報を踏まえた上でリソースの配分などを検討しなければならない．しかし，AI ワークロードを走行させるハードウェアリソースの有効活用･運用の制御手法は未だ確立されてない．そこで，ワークロード毎にサーバ性能の自動チューニングを行う機械学習向けのコンピュータシステムの構築を考えたい．

本研究ではコンピュータシステムの構築に向けて，AI ワークロードに特化したコンピュータシステムの最適リソース設計を目的として，機械学習系アプリケーションのパフォーマンス測定のためのベンチマークである MLPerf[3] を用いた AI ワークロードの比較及び特徴分析を行う．

2　関連技術

Facebook 社では保有するデータの大部分を機械学習パイプラインに流しており，GPU と CPU の両方を活用するトレーニングと CPU を活用するリアルタイム推論でシステムを使い分けている．また，リアルタイム推論では入力データが大きいことから必要リソースも異なってくるなど，文献 [4] では機械学習に基づく AI ワークロード挙動の特徴分析の重要性が問われている．

ボトルネックを分析し，リソース配分を行う自動ワークロード管理機能や接続リクエストを均等に分散する機能を備えた CPU やサーバは既に開発されている．CPU コアやクラスタ単位で負荷に応じて電圧と動作周波数を切り替える Dynamic Voltage and Frequncy Scaling を実行する Intel の “Haswell”[5] 以降の CPU アーキテクチャ，CPU アーキテクチャや多くの GPU を積んで AI ワークロードに特化させた NVIDIA のクラウドサーバプラットフォーム “HGX-2”[6] などのハードウェア群が例として挙げられる．

3　実験

本研究では，AI アプリケーションの実行時のハードウェア情報の分析を目的として，MLPerf を利用し各ベンチマークの性能評価，比較を行う．MLPerf には画像分類の image classification，物体認識の single stage detector，object detection，自然言語処理の recommendation，sentiment analysis，翻訳の rnn translator，translation，音声認識の speech recognition，強化学習の reinforcement（図表では IC，SSD，OD，RM，SA，RT，TL，SR，RI と略記する）といった 9 つのベンチマークが実装されている．情報取得には Zabbix[7] を用いる．Zabbix は Zabbix 社が開発しているネットワーク管理ソフトウェアで，本研究では perf，nvidia-smi のコマンドで取得できるプロセッサやメモリに関する情報を分析対象とする．ベンチマーク毎の特徴分析のために，これらのコマンドによって得られる情報をベンチマーク実行時に取得する．情報取得は 1 分間隔で行い，今回は CPU や GPU の利用率，メモリの利用量に注目して分析を行った．実験環境を表 1 に示す．

表 1: 実験環境

OS	ubuntsu 16.04
サーバ	FUJITSU Primergy RX2540 M4
CPU	Intel Xeon Skylake 2 ソケット x20 コア 2.4GHz Gold 6148 150W
GPU	NVIDIA Tesla V100 16GB
Storage M2.SSD	290GB read 0.87GB/s，write 1.75GB/s
Memory	192GB DDR4 2666MHz
Python	3.50
CUDA	9.2

4　実験結果

各ベンチマーク実行時の GPU/CPU 比率を図 1，プロセッサの平均利用率を表 2，メモリ最大利用量を図 2，GPU を変更して MLPerf を実行した際に要したジョブ時間の比較結果を表 3 に示す．

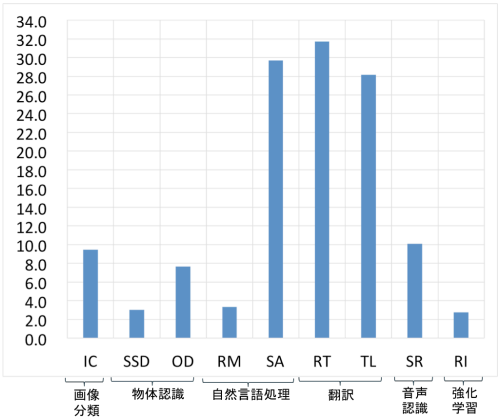

図 1: GPU/CPU 平均比率

図 1 の縦軸は GPU 利用率の平均値を CPU 利用率の平均値で割ったものを表している．本研究で利用したサーバは 2 ソケットなので CPU1 ソケット当たりの利用率に換算している．GPU/CPU 比率を CPU1 つ当たりの GPU の必要量の概算に用いる．表 2 は CPU と

表 2: CPU と GPU の平均利用率

ベンチマーク	CPU 利用率 (%)	GPU 利用率 (%)
IC	5.1	95.4
SSD	9.9	59.8
OD	4.9	74.5
RM	6.7	44.3
SA	1.4	82.0
RT	1.5	95.5
TL	1.5	83.9
SR	3.2	65.1
RI	11.4	62.7

GPU の平均利用率を表している．表 2 を見るとアプリケーション全体としては GPU ネックの傾向にあり，翻訳系アプリケーションでは特に GPU の必要量が大きいことが確認できる．また，系統の異なるアプリケーションではプロセッサの利用率に大きく違いがある．

表 3 は GPU を V100 から P100 に変更し，それぞれで各ベンチマークを実行した際のジョブ時間の比較結果である．表 2 の GPU 平均利用率と比べると，GPU

表 3: ジョブ時間の比較 - GPU

ベンチマーク	P100	V100	P100 / V100
IC	4:33:04	3:01:13	1.51
SSD	3:12:18	3:09:12	1.02
OD	2:57:51	2:23:26	1.24
RM	1:09:07	1:09:34	0.99
SA	1:48:14	1:22:50	1.31
RT	4:05:28	2:48:18	1.46
TL	4:29:21	2:59:55	1.50
SR	15:07:34	10:53:32	1.39
RI	6:08:37	5:46:00	1.07

平均利用率が高いベンチマークほど GPU を変更した際のジョブ実行時間の変化が大きくなっている．

これらの結果から GPU 利用率が高いジョブは GPU 性能によるジョブ性能向上効果が高く，CPU 性能によるジョブ性能の差は小さいと推測できるため，AI の種類によって最適な GPU/CPU リソース設計ができる．

図 2 の縦軸はメモリ利用量の最大値を表している．GPU メモリ利用量と比べ，殆どのアプリケーションでは CPU メモリ利用量が大きくなっている．これはアプリケーションが要求するデータサイズに対して GPU メモリ容量が不足しているためである．プロセッサ利用率の結果も踏まえると，容量不足の問題が解決できればより効率良いアプリケーションの利用が可能になると推測できる．画像分類系アプリケーションや翻訳系アプリケーションはどちらも GPU 利用率の高いワークロードだが，CPU メモリ利用量に大きな差がある．

5 まとめと今後の課題

GPU を用いた AI アプリケーション群でも GPU の必要量に大きく差があるが，全体的には GPU ネックの

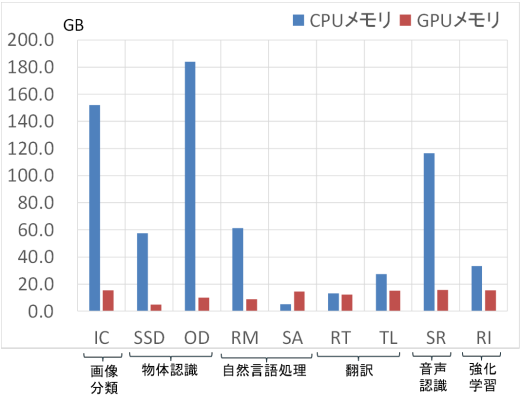

図 2: メモリ最大利用量

傾向が見られた．GPU 利用率が高いジョブは GPU 性能によるジョブ性能向上効果が高く，CPU 性能によるジョブ性能の差は小さいと推測した．また，メモリ利用量はアプリケーションの分野によって大きく異なっており，メモリも重要であることが分かった．

今後は I/O や周波数，メモリアクセスに注目した特徴分析や，異なるサーバ機種で MLPerf を実行した際のデータ比較も行いたい．

謝辞

本研究の一部はお茶の水女子大学と富士通研究所との共同研究契約に基づくものである．また，本研究にご協力頂いた富士通コンピュータテクノロジーズの鈴木孝規氏に深謝する．

参考文献

[1] Camilo Mora, Randi L Rollins, Katie Taladay, Michael B Kantar, Mason K Chock, Mio Shimada, and Erik C Franklin. Bitcoin emissions alone could push global warming above 2°c, 2018.

[2] Zhenhua Liu, Yuan Chen, Cullen Bash, Adam Wierman, Daniel Gmach, Zhikui Wang, Manish Marwah, and Chris Hyser. Renewable and cooling aware workload management for sustainable data centers, 2012.

[3] MLPerf v0.5. https://mlperf.org/. Accessed: 2018-12.

[4] Kim Hazelwood, et. al. Applied Machine Learning at Facebook: A Datacenter Infrastructure Perspective. 2018 IEEE International Symposium on High Performance Computer Architecture (HPCA), pp. 620-629. IEEE (2018).

[5] Haswell. https://ark.intel.com/ja/products/codename/42174/Haswell. Accessed: 2018-12.

[6] HGX-2. https://www.nvidia.com/en-us/data-center/hgx/. Accessed: 2018-12.

[7] Zabbix. http://www.zabbix.com/. Accessed: 2018-12.