

仮想環境における Hadoop 処理時間予測に関する実験と考察 Experimental and Consideration on Predicting Hadoop Processing Time in Virtual Machine Environments

古畑 俊一郎[†] 祖父江 恒夫[†] 菅原 秀幸[‡]
Shunichiro Furuhata Tsuneo Sobue Hideyuki Sugawara

1. はじめに

近年、走行している車両からデータを集めて解析するプローブカーシステムが注目されている。走行車両のデータを収集して解析することで、渋滞情報やエコ運転支援といったプローブカーサービスが実現できるためである。これらのプローブカーサービスを実現するため、走行車両は常にデータを取得してセンタに送信する。そのためサービス対象車両が増加すると、解析するデータ量も増加することとなる。そこで、プローブカーサービスを提供するサービス事業者には、(1) 解析処理に必要な計算機リソースのスケールアウト、及び(2) その解析処理の最小コストでの実現、という2つの要求がある。

本稿では、計算機リソースのスケールアウトを Hadoop による並列処理で実現することを想定する。つまり、データ量の増減に応じて並列処理台数を増減させる。ここで、サービス事業者がプローブカーサービスを提供することを考えると、サービスが提供する情報更新頻度は車両の増加に伴い低下してはならず、対象車両が増加しても一定時間で情報を更新しなければならない。そのため、入力データ量に応じて並列処理台数を変更することで、処理時間を一定にすることが必要となる。

一方、計算機リソースを柔軟に変更するための手段として仮想サーバを用いたクラウドサービスがあげられる。仮想サーバを用いることで計算機リソースの柔軟な調整ができるためである。そのため、上記(1)(2)の目的を達成するためには、仮想サーバで Hadoop を動作させる際の、入力データ量と並列処理台数に対する処理時間予測が重要である。本稿では、仮想サーバを用いた Hadoop 処理時間予測を実現するため、先行研究による処理時間予測手法[3]の適用条件について考察する。

2. 従来手法

Hadoop での先行研究による処理時間予測手法[3]は、Hadoop の処理時間が入力データ量と比例関係にあることを利用している。ある Hadoop 環境で少量のデータで処理時間を計り、予測式のパラメータとして当てはめる。この予測式の項にデータ量を入力することで、その Hadoop 環境での処理時間を予測することができる。

しかし、仮想環境では仮想デバイスを複数の仮想サーバで共有するため、物理環境では発生しないデバイスの共有による I/O 帯域の減少が起こる。帯域の減少により I/O 待ちが発生し、それに伴う処理遅延により物理環境とは異なった処理時間特性を持つ。そのため、従来手法の予測式が適用できない。

[†] 株式会社 日立製作所

[‡] 株式会社 日立ソリューションズ

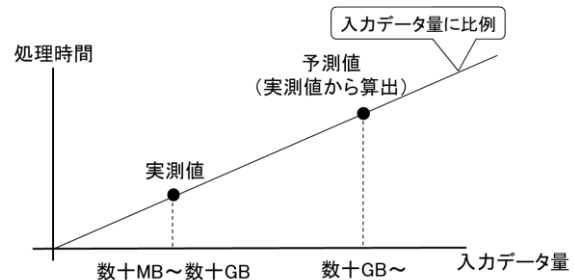


図1 処理時間予測手法の概要

3. 仮想環境への従来手法の適用検証

3.1 実験概要

従来手法を仮想環境に適用した結果を確認するため、実験し結果を比較した。実験は仮想環境に構築した Hadoop クラスタを利用した。

仮想化機構が仮想化するデバイスとして、CPU、メモリ、ディスク I/O、ネットワーク I/O がある。本実験ではディスク I/O の帯域低下の影響を確認するための実験を行うこととする。そこで、CPU は仮想サーバ 1 台に 1 コアを専有させる、メモリは各仮想サーバに固定の領域を割り当てる、ネットワークは物理環境内で完結させるといった環境を構築し速度低下やネットワーク I/O の帯域低下が起こらないように仮想環境を構築した。処理時間計測のプログラムは、ディスク I/O に大きな負荷がかかるアルゴリズムである Sort [4]を利用した。

3.2 実験結果

物理環境と仮想環境で Sort を実行し比較した結果を図 2 に示す。物理環境に比べて仮想環境の方が、小さな入力

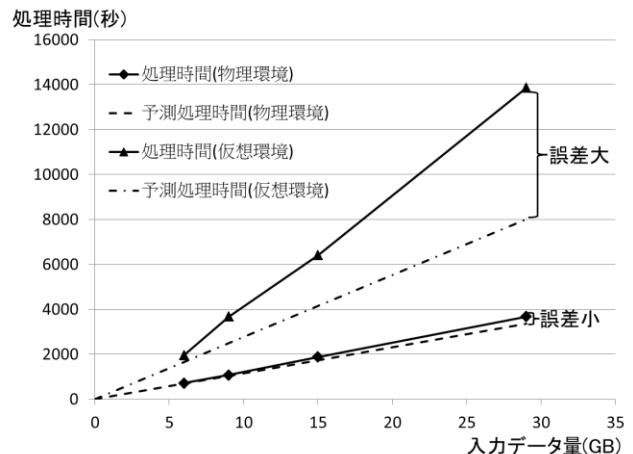


図2 仮想環境と物理環境の Sort 処理時間・予測処理時間

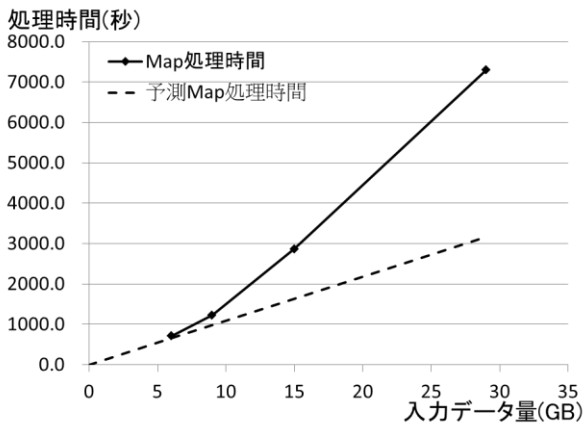


図3 仮想環境での Map 処理時間・予測時間

データ量に対して大きく予想処理時間に誤差が発生している。具体的には、処理時間に対して 9GB では物理環境で 3.0%, 仮想環境で 14.4%, 誤差があるのに対して、29GB では、物理環境で 9.0%, 仮想環境で 42.2%の誤差となっている。

4. 予測時間誤差の考察

実験結果より、仮想環境では物理環境に比べて誤差が大きくなることがわかった。そこで Hadoop の処理フローに注目し誤差の原因を考察した。

まず、仮想環境での Map 処理, Shuffle 処理, Reduce 処理の処理時間と予測処理時間を確認した。その結果, Map 処理の処理時間が予想処理時間に対して大きな誤差が発生していることがわかった(図 2)。

Map 処理に大きな誤差が発生する理由は、仮想化によりディスク I/O 帯域が狭まった環境で Map 処理での大量のディスクアクセスが処理待ちを発生するためである。Map 処理は Map 本体処理と Map 後処理の 2つのフェーズで構成されている。Map 本体処理は Map プログラムに従って入力したデータを分割し処理する。処理した結果は、バッファ領域に格納されバッファが足りなくなると中間ファイルとしてディスクに書き出す。中間ファイルは Map 後処理に渡される。Map 後処理は、渡されたデータをソートする。そしてソートされた複数の中間ファイルをマージする。この時、マージのための中間ファイルの読み込み、マージ結果の書き込みのため大量のディスクアクセスが発生する。仮想化によりディスク I/O 帯域が狭くなっているため、処理待ちが起こる。処理待ちの発生を考慮していない従来手法による予測では Map 処理の予測処理時間に誤差が発生する。それに伴い、処理全体の予測処理時間にも誤差が発生する。

5. 仮想環境で Hadoop を使用する場合の指針

実験結果及び考察より、仮想環境では Map 処理でのディスクアクセスが処理時間予測の誤差の原因になるとわかった。そこで、誤差を避けるための指針として、仮想環境では Map 本体処理が生成する中間ファイルの量が少ない処理のみを実行し仮想化による影響を避ける、Map と Reduce 処理の役割分担を変え中間ファイルの量を削減するように MapReduce プログラムを設計するといった対策をとることで従来手法により処理時間予測が可能になる。また、本稿では言及しなかったが、Map 後処理に関

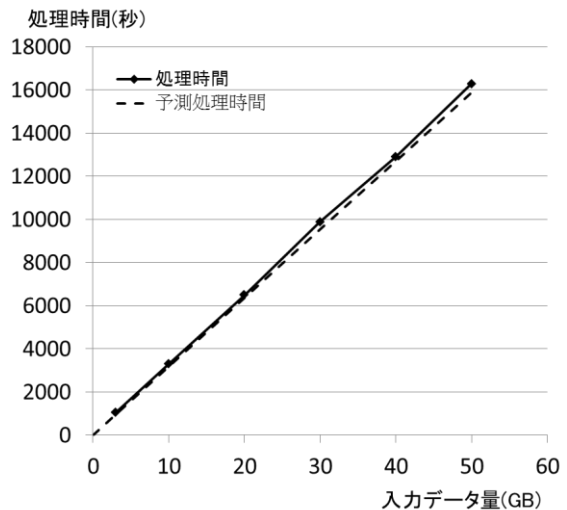


図4 提案指針に沿ったプログラムの処理結果

わる Hadoop パラメータをチューニングすることにより、従来手法による予測がより適切なものになると考えられる。以上を仮想環境での Hadoop 処理時間を予測する際に従来手法を適用する指針とする。

なお指針の有効性を確認するため、指針に沿って設計した Map プログラムを含む処理を実施し、有効性を確認した。その結果、提案指針に沿って MapReduce プログラムを設計することで、仮想環境でも従来手法の適用が可能となることが確認できた(図 4)。

6. おわりに

クラウドサービスでプローブカーサービスを実現することを想定し、仮想サーバ上での Hadoop 処理時間予測について考察した。仮想化によりディスクアクセスがボトルネックになる場合に従来手法の予測値からのずれが生じると考え、実験によって検証した。その結果、Map 処理後のディスクアクセス回数が多い処理では従来手法が適用できないことがわかった。また、ディスクアクセスが少ない場合は、従来手法が利用可能であるとの見通しを得た。

今後は、処理時間予測における Hadoop の設定パラメータとタスク分割数によって、処理時間予測に対する誤差がどう生じるかを測定し、実際の運用環境で処理時間予測が可能となる設計指針の策定をめざす。

参考文献

- [1]Hadoop <http://hadoop.apache.org/>
- [2]NIST, The NIST Definition of Cloud Computing, <http://www.nist.gov/itl/cloud/upload/cloud-def-v15.pdf>
- [3]NTTデータ, “平成 21 年度産学連携ソフトウェア工学実践事業 事業 成果 報告 書 ”, http://www.meti.go.jp/policy/mono_info_service/joho/downloadfiles/2010software_research/clou_dist_software.pdf
- [4]S. Huang, J. Huang, J. Dai, T. Xie, and B. Huang, “The HiBench benchmark suite: Characterization of the MapReduce-based data analysis,” ICDEW 2010.