

クラウドコンピューティングにおけるコストを考慮したインスタンス分類手法 Classification of Instances using Cost Information in Cloud

伊藤 耕平[†] 小坂 隆浩[†]
Kohei Ito Koita Takahiro

1. はじめに

近年、クラウドコンピューティングが普及し、ユーザは非常に手軽に大規模な計算資源を利用することが可能になった。Amazon が提供するクラウドコンピューティングサービスである Amazon EC2(Amazon Elastic Computing Service)[1]の形態は IaaS と呼ばれ、ユーザが料金を支払うことで、インターネットを通じてインスタンスと呼ばれる仮想計算機を利用することができるようになる。インスタンスは複数の種類が用意されており、それぞれ CPU やメモリなどの性能と 1 時間あたりの料金が異なる。そのため、同様のアプリケーションの使用を想定した場合でも、選択したインスタンスによって処理時間や料金が異なる場合がある。ユーザが処理にかけることのできる資金は無限ではないため、できるだけ料金を軽減させることが望ましい。料金を軽減させるためには、アプリケーションに応じた適切なインスタンスを選択する必要があるが、インスタンスの性能値や料金についての比較は可能だが、処理の結果や利用料金については実際にインスタンスを利用してみなければ比較することはできないため、ユーザのインスタンスの選択は困難である。適切なインスタンス選択ができない場合には、必要以上の料金を消費してしまうこともある。

本研究では、AmazonEC2 における利用料金の軽減を目的とし、インスタンスの分類手法の提案を検討する。

2. Amazon EC2

2.1 概要

AmazonEC2 は、インターネットを通じて、仮想計算機を提供するクラウドコンピューティングサービスである。提供される仮想計算機をインスタンスと呼び、インスタンスには複数の種類が用意されている。インスタンスのような計算機資源が実際に存在しているデータセンタをリージョンと呼び、ユーザはリージョンを自由に選択することができる。インスタンスはハードウェア資源とソフトウェア資源をひとまとまりにしたものであり、ユーザはインスタンス上で自由に環境を構築できる。インスタンスはそれぞれ、性能値と 1 時間あたりの利用料金が設定されており、ユーザはどのインスタンスを利用するかを選択する。AmazonEC2 の各インスタンスの詳細を表 1 に示す。ここでのインスタンスの例は、リージョンが北カリフォルニア、OS が WindowsServer2008 の場合である。各インスタンスの CPU 性能は ECU という値で表され、1ECU は 1.0-1.2GHz 2007 Opteron, または 2007 Xenon プロセッサの CPU と同等の処理能力である。

インスタンス	ECU	コア数	1コアの ECU	メモリ (GB)	I/O性能	ストレージ (GB)	1時間あたりの利用料(\$/h)
m1.small	1	1	1	1.7	標準	160	0.13
m1.large	4	2	2	7.5	高速	850	0.52
m1.xlarge	8	4	2	15	高速	1690	1.04
t1.micro	最大2	1	最大2	0.613	高速	従量制	0.035
m2.xlarge	6.5	2	3.25	17.1	標準	420	0.69
m2.2xlarge	13	4	3.25	34.2	標準	850	1.38
m2.4xlarge	26	8	3.25	68.4	高速	1690	2.76
c1.medium	5	2	2.5	1.7	標準	320	0.31
c1.xlarge	20	8	2.5	7	高速	1690	1.24

表 1 各インスタンスの詳細

2.2 Amazon EC2 利用における問題点

Amazon EC2 では複数のインスタンスが提供されており、ユーザは自由にインスタンスを選択することができる。しかし、インスタンスの選択にあたって、ユーザが判断する材料となるのは Amazon が示すインスタンスの性能値と、1 時間あたりの利用料金だけである。例えば、m1.small と c1.medium の 2 つのインスタンスを比較した場合、1 時間あたりの利用料金は約 2.4 倍であるが、CPU 性能の差は 5 倍である。したがって、CPU 性能だけに依存する処理を行った場合、m1.small の処理時間は c1.medium の 5 倍かかる。よって、処理時間、利用料金ともに c1.mediumの方が低くなる。このように、CPU 性能だけに依存するアプリケーションや処理であれば、上記のように性能値と 1 時間あたりの利用料金だけで、適切なインスタンスを選択することができる。しかし実際のアプリケーションは CPU 性能だけでなく、メモリ性能は I/O 性能など、複数の処理性能の影響を受ける。そのため、性能値と 1 時間あたりの利用料金だけでは、利用したいアプリケーションに対して、どのインスタンスが適切であるか判断することは難しい。インスタンスの選択によっては、必要以上に利用料金を消費してしまう場合もある。

3. インスタンス分類手法

Amazon EC2 の利用料金を軽減させるためには、ユーザが利用するアプリケーションに応じて、適切なインスタンスを選択できればよい。そのためには、アプリケーションに応じたインスタンス分類手法を検討することが必要である。インスタンス分類手法とは、アプリケーションに対応した適切なインスタンスを選択するための指標のことである。各インスタンスに対してアプリケーションを実行し、処理時間や処理性能、利用料金を評価する。評価から、適切なインスタンスを選択できる条件を検討し、インスタンス分類手法を導く(図 1)。この分類手法に基づいてインスタンスを選択することで、利用料金の軽減を目指す。

[†] 同志社大学 工学部 情報システムデザイン学科

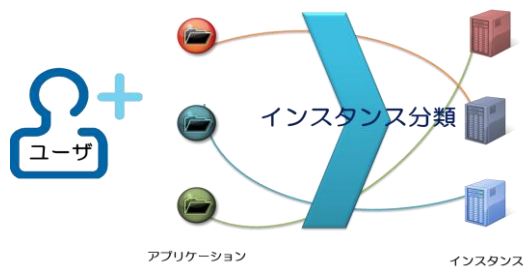


図1 インスタンス分類手法のイメージ

4. 評価実験

4.1 実験概要

インスタンス分類手法の検討に向けて、インスタンスの評価実験を行う。複数のアプリケーションを用意し、各インスタンスに対してアプリケーションそれぞれを実行する。実行中のインスタンスの CPU 使用率やメモリ使用量、I/O 使用量などの処理性能やアプリケーションの処理時間、処理で生じた利用料金を評価する。評価のイメージを図2に示す。

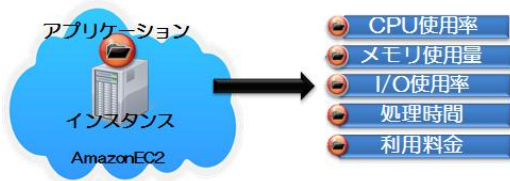


図2 評価イメージ

4.2 実験環境

インスタンスの評価実験の環境として、並列分散処理プラットフォームの一つである Hadoop を用いる。Hadoop は大規模なデータ処理に適した技術であり、今回のインスタンス評価実験で扱うアプリケーションでは、ある程度大規模なデータ量を想定する。Hadoop は、クラスタ全体を管理するマスターサーバと、実際に処理を行うスレーブサーバと呼ばれるサーバ群から成る。Hadoop は大きく分けて、クラスタ環境で並列分散処理を行うためのフレームワークである MapReduce と、MapReduce で処理するデータを扱う分散ストレージである HDFS で構成されており、これらが動作することで一つの Hadoop プラットフォームとなる。Hadoop のサーバ構成図を図3に示す。

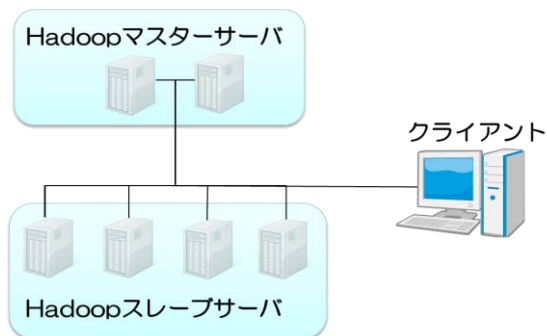


図3 Hadoop サーバ構成

Hadoop にはプラットフォームに関するパラメータと、アプリケーションに関するパラメータが存在し、パラメータの値によって MapReduce と HDFS の挙動が決めるため、パラメータチューニングの問題が挙げられるが、今回は大規模分散処理基盤としての利用のために Hadoop を用いるため、パラメータに関しては考えない。

Hadoop のサーバ構築に Amazon EC2 のインスタンスを用いる。OS は Windows Server 2008 を使用して実装する。構築した Hadoop 環境上でアプリケーションを実行し、評価を行う。

5. 評価結果

評価実験に基づいて、アプリケーションごとに、各インスタンスについて評価を行う。インスタンス評価の一例として、ルート計算アプリケーション、行列演算処理アプリケーション、ファイル読み込みアプリケーションの料金に関する評価結果を図4に示す。

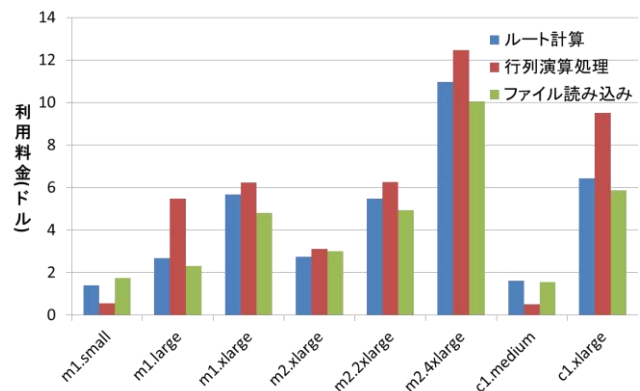


図4 評価結果

6. まとめと今後の展望

本研究では、Amazon EC2 におけるコストを考慮したインスタンスの分類手法について述べた。ユーザにとって、インスタンスの性能値と一時間あたりの料金だけではインスタンスの選択が困難である。そのため、インスタンスの分類手法に基づいた適切なインスタンス選択が必要である。今回は、一例としてしかインスタンスの評価を行えなかったため、今後は Hadoop 環境において、各インスタンスについて処理性能や処理時間、料金を評価し、その結果に基づいてインスタンスの分類手法を検討していく必要がある。

参考文献

- [1] Amazon Elastic Compute Cloud (Amazon EC2), amazon web service, (2008).