

インタークラウド環境における Web 三層モデルと CDN を 組み合わせた多目的資源割当最適化に関する検討

川勝崇史^{†1} 棟朝雅晴^{†2}

Web 三層モデルのシステムとキャッシュサーバーを用いて負荷分散を行う CDN を組み合わせたシステムのマッピングについて最適化を行う。またこのシステムにインタークラウド環境を用いてサーバーのスケジューリングを行ないエネルギーの効率化を行う。本研究ではコストや CO2 排出量の削減、レスポンスを最適化することを目的としてインタークラウド環境における web 三層モデルと CDN を組み合わせたシステムについて三つの要素を多目的遺伝的アルゴリズムを使って同時に資源割当最適化を行い実験の妥当性・有効性について検討する。

Multi-purpose optimization of resource allocation combined CDN and Web three-layer model in inter-cloud environment

Takafumi Kawakatsu^{†1} Masaharu Munetomo^{†2}

This paper proposes a framework of optimizing allocations of the virtualized systems consisting of virtual machines based on web three-layer model and Contents Distribute Network (CDN) to the inter-cloud system, in order to maximize energy efficiency and also minimize carbon emissions. In this study, we solve a multi-objective optimization problem with a multi-objective genetic algorithm (MOGA) to optimize cost, performances and carbon emissions at the same time, and also investigate the effectiveness of the proposed framework.

1. 研究の背景と目的

近年 ICT の急速な進歩に伴って、サーバーやネットワーク、データセンターによる電力消費が増加している。これは化石燃料で電力から電気を生成するので温暖化に繋がる。電力のコストやネットワークのレスポンスだけでなく二酸化炭素排出量についても考える必要がでてきた。

そこでインタークラウドを用いて各データセンターのクラウドを相互に接続すると異なるデータセンター間でリソースを相互利用できるのが各データセンターではコストや二酸化炭素排出量が同じではないためどのデータセンターの資源を割り当てるかが重要になってくる。

本研究では Web 三層モデルと CDN(Contents Distribute Network)を組み合わせたシステムをどのインタークラウド環境にマッピングするかを最適化する。具体的にはコスト・レスポンス・二酸化炭素排出量の削減という三要素を同時に多目的最適化するモデルを検討する。

2. 数理モデル

本実験はWWW層とAPP層とDB層の三層で形成される静的なWeb三層モデルとその上にCDNを組み合わせたシステムにVM(Virtual Machine)を割り当てるモデルを考える。CDNはReverse Proxy Serverと同じくキャッシュを保存しておき新しいリクエストがあった場合にはそのリクエ

ストをキャッシュとして保存する

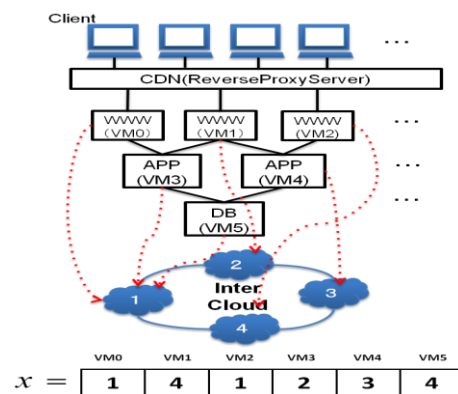


図 1 CDN と Web 三層のモデル

この時、CDNはAmazon Cloud frontを例として各地域に配置し料金もこれにならって設定する。以上を踏まえた上でコスト・レスポンス・二酸化炭素排出量を最適化する。

また各個体表現についてはnを個体数、Cをクラウドの個数として以下のように表現する。

$$x = \langle x_1, x_2, \dots, x_n \rangle \quad (1)$$

各 x_i をVMのIDとすると $x_i \in \{0, 1, 2, \dots, C\}$ ($i = 0, 1, \dots, n$)は各VMがどこのクラウドを使用しているかをベクトルで表す。目的関数Fは以下で定義する。

$$F(x) = \{f_{\text{cost}}(x), f_{\text{response}}(x), f_{\text{carbonemission}}(x)\} \quad (2)$$

電気使用量 E_j について以下の(3)式で定義する。

$$E_j = E_s + E_a \times E_f + \alpha \quad (3)$$

^{†1} 北海道大学大学院情報科学研究科
 Graduate School of Information Science and Technology, Hokkaido University.

^{†2} 北海道大学情報基盤センター
 Information Initiative Center, Hokkaido University

ただし、クラウドのサイト数をN、各クラウドの番号をj、全電気使用量をE_j、各VMの待機電力をEs、サーバー使用率をEa、サーバー100%使用時の電力Ef、ネットワーク上の電気使用量α。

次に電気使用料のコスト関数 f_{cost}(x)について以下の(4)式で定義する。

$$f_{cost}(\vec{x}_i) = \sum_j^N (CP_j + rCost_j + tCost_j + cCost_j) \quad (4)$$

ただし、各クラウドにおける 1kw 当たりの電気使用料金を CP_j、VM リソース・データ転送・CDN 使用した料金をそれぞれ rCost_j、tCost_j、cCost_j。

レスポンスのコスト関数 f_{response}(x_i)について以下の(5)式で定義する。

$$f_{response}(\vec{x}_i) = Rp + Rd(m, n) \quad (5)$$

$$Rp = t_i/p \quad (6)$$

$$Rd(m, n) = (\beta / c') \times 2(s) \quad (7)$$

ただし、処理時間を Rp、処理量を t_i、通信速度を p、応答時間を Rd(m, n) ただし m, n は通信する VM 同士の位置とする。実測距離を β、光の速度 c'、このとき(7)式に定義するように応答の場合には上り下りとあわせて二倍となる。

炭素排出量のコスト関数 f_{carbonemission}(x_i)について以下の(8)式で定義する。

$$f_{carbonemission}(\vec{x}_i) = \sum_{j=0}^N \rho_j E_j \quad (8)$$

ただし、各クラウドにおける 1kw 当たりの炭素排出率を ρ_j

3. 多目的最適化

多目的最適化とは目的関数が複数存在する最適化のことである。実際に目的関数が単一である問題もありますが複数の問題を最適化する必要がある問題も存在する。今回は目的関数が三つあるので多目的最適化を用いる。

実験で使用したアルゴリズムは多目的遺伝的アルゴリズム NSGA-II である。特徴は非優越ソートによるランキング方法、混雑距離の導入、混雑度トーナメント選択の導入の三つである。各 VM における消費電力について Es、Ef は実際の電力量をもとに設定する。また、各国の電気料金は今回ウェブサービスに焦点をおいているので産業用の電気料金で計算する。各料金コストについて、rCost と tCost をクラウド内で規模を自在に変更可能なウェブサービスである Amazon Elastic Compute Cloud (Amazon EC2) を元に各料金コストを設定する。また cCost については Amazon Cloud Front をもとに各料金コストを計算する。各炭素排出量については主に二酸化炭素を排出して発電する石炭・石油・ガスの割合と 1Kw/h の排出量によって計算し原子力や水力その他はコスト計算には使わないものと

する。CDN を考えない場合のデータを図 2 に示す

4. まとめと今後の課題

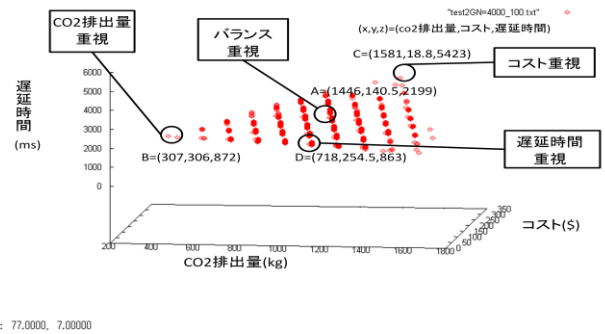


図 2 実験結果

WEB 三層モデルを固定し CDN を考えない場合について、まず実際の値と照らし合わせながらコスト関数を作った点について、電力消費にかかわる二酸化炭素排出量と料金コストはおおむね実測値とほぼ同じであるので上手くいった。VM の処理速度は実測値で行ったので処理時間は計算出来るが応答時間は理論値で考えた。この点に関しては、レスポンスのコスト関数を理論値(光の速度と実測距離)ではなく、ルーターやネットワーク上のホップ数で近さ遠さを加味してより現実に計算する方法を考える必要があった。次に、NSGA-II アルゴリズムについて今回の実験結果から環境がより複雑になりより難しい状況で最適化を行うにあたり容易に目的関数を増やせ各目的関数の優良個体を保存しながら最適化出来使用者が自分の望む最適化を選ぶというのは利点である。

また CDN を含めて考えた場合、各 VM についてあまり使われていないクラウドやコストの低いクラウド、二酸化炭素排出量の少ないクラウドなどを切り替えることで最適化を行うと CDN により遅延時間は短くなるが、逆にコストは上がると予想される。本研究は次の課題として CDN を含めた場合の検討があっているのか考えていく。最終的な目的としては各データセンターで作るクラウド繋ぎ自動で資源割当のスケジューリングをできるように考えている。

参考文献

- 1) 棟朝雅晴, 遺伝的アルゴリズム—その理論と先端的手法—, 森北出版, 2008.
- 2) Kalyanmoy Deb, Multi-Objective Optimization using Evolutionary Algorithms, 1995.
- 3) Bang Minyoung, Asim Munawar, Omar Abdul-Rahman Masaharu Munetomo, Kiyoshi Akama, Optimal Resource Allocation for Distributed Cloud Systems, 2011.
- 4) Rajkumar Buyya, An-ton Beloglazov, Jemal Abawajy, Energy-Efficient Management of Data Center Resources for Cloud Computing: A Vision, Architectural Elements, and Open Challenges, 2010.
- 5) IEA, "Energy Balances of OECD Countries 2010", "Energy Balances of non-OECD Countries 2010" OECD/IEA, Energy Prices and Taxes, Volume 1999-1/Volume 2005-1/Volume 2011-1, 2
- 6) <http://aws.amazon.com/jp/>, <http://aws.amazon.com/jp/cloudfront/>