

5J-7 消費電力量を含むコストを考慮したデータインテンシブアプリケーションのハイブリッドクラウドにおける負荷分散ミドルウェアの提案

笠江 優美子[†]

小口 正人[†]

[†]お茶の水女子大学

1. はじめに

近年、コンピュータシステムにおける情報量が爆発的に増大しており、そのような大量のデータを効率よく処理するシステムが求められている。そこで本研究では、そのようなシステムをハイブリッドクラウド環境において実現するミドルウェアを提案する。また、近年の世界的なエコ志向により、クラウド構築側が抱えるコンピュータ機器の消費電力量の増大が問題となっている。本ミドルウェアは、ハイブリッドクラウド環境において、消費電力量を含む金銭的成本を抑える負荷分散を実現する。本ミドルウェアによって、時間的成本を重視した負荷分散や、消費電力量を含む金銭的成本を抑えることを重視した負荷分散を、ユーザの指示によるパラメータ設定で実現する。

本稿では、提案するミドルウェア実行時のジョブの処理時間、パブリッククラウドの従量制料金、プライベートクラウドの消費電力量料金を測定し、それらを用いて本ミドルウェアの評価を行った。

2. 提案する負荷分散ミドルウェア

本ミドルウェアの特徴は2つある。1つ目の特徴は、実行環境としてハイブリッドクラウドを想定している点である。ハイブリッドクラウド環境での負荷分散を考えた場合、ジョブの処理時間とパブリッククラウドの従量制料金のコストバランスが求められる。本提案では、リソースを使い切った後、負荷分散を行うことでコストを抑えていき、効率よくジョブを処理することを目指す。

2つ目の特徴として、データインテンシブアプリケーションをジョブの対象としている点である。データインテンシブアプリケーションの場合、処理がI/O待ちとなっている場合が多く、CPU負荷から適切な判断ができないため、本ミドルウェアではDisk I/Oを用いる。図1はデータインテンシブなジョブを実行させた時のDisk I/Oと処理時間の推移である。ジョブの投入量が増えると、Disk I/Oの値は飽和状態となり、この状態になると飽和していない状態を表したbaselineに比べ、実行時間が長くなっている。このことから、本ミドルウェアでは、リソースが飽和したと判断されると負荷分散を行う。

上記2つの特徴を実現するため、ミドルウェアは定期的にDisk I/Oを測定し、飽和状態のDisk I/Oの値(S値)を任意回上回ったら(U値)、別のインスタンスへ負荷分散を行い、S値を任意回下回ったら(L値)、そのインスタンス

は飽和していないと判断する。さらに、ユーザが飽和判断レベルの設定を行うパラメータを変化させることで、ジョブの処理時間や消費電力料金を含む金銭的成本を重視した負荷分散を行うことを選択できる。この飽和判断レベルのパラメータを変化させるということは、U値とL値を変化させることを意味する。

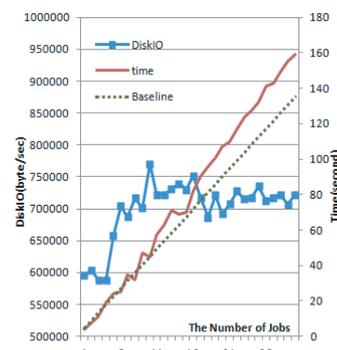


図 1: データインテンシブなジョブ実行時の Disk I/O と処理時間

3. ミドルウェアの実行実験

3.1 実験概要

本ミドルウェアを実行させる際、実験では図2に示すようなハイブリッドクラウド環境を構築し、8台のインスタンスに対し負荷分散を行なった。

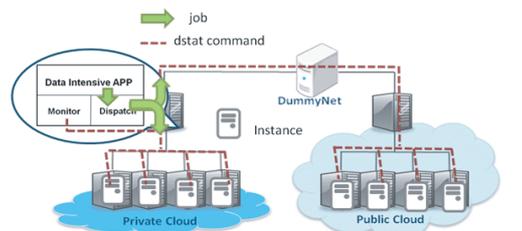


図 2: ハイブリッドクラウド環境

Level	U値	L値
I	2	5
II	3	5
III	4	5
IV	5	5
V	6	5
VI	7	5
VII	8	5
VIII	9	5
IX	10	5

図 3: Level

実験は、図3のように飽和判断レベルを変化させた。この飽和判断レベルでは、小さいほど負荷分散されやすく、大きいほど負荷分散されづらいことを意味する。

本実験で投入するジョブは、データインテンシブなアプリケーションを想定するため、PostgreSQLのベンチマークであるpgbenchを用いた。実験ではジョブを2秒間隔で計200回投入した。1ジョブあたり処理時間は平均9秒程度である。実験では、上記のレベルを変化させて実行させたときの、ジョブの実行時間と、パブリッククラウドの従量制料金、プライベートクラウドの消費電力料金を測定した。ここで、測定対象にパブリッククラウドの消費電力料金がわからないのは、実際にパブリッククラウドの消費電力料金がわからないことは想定しづらく、わかったとしてもユーザのモチベーションに依存するものであり、従量制料金に消費電力料金が含まれると考えているためである。

Proposing the Method for Data Intensive Application's Load Distribution Considering the Costs Balances including Power Consumption Cost on Hybrid Cloud

[†] Yumiko Kasae, Masato Oguchi
Ochanomizu University (†)

3.2 実行結果

図4, 5, 6は, ミドルウェア実行時のジョブの処理時間と, パブリッククラウドの従量制料金, プライベートクラウドの消費電力量の測定結果である.

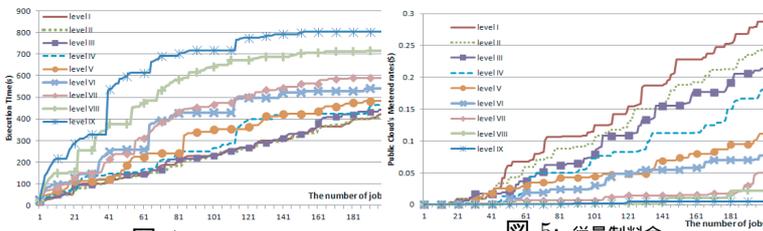


図 4: 処理時間

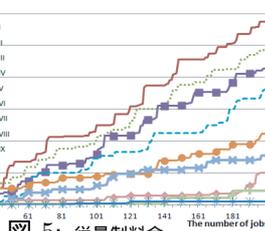


図 5: 従量制料金

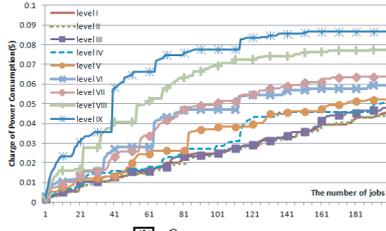


図 6: 消費電力量

図5は, パブリッククラウドに投入したジョブの実行時間, 借りたインスタンス数, 従量制料金単価を掛けたものである. パブリッククラウドの従量制料金単価は, Amazon EC2の従量制料金を元に1時間あたり\$0.5として計算した. 図6では, 東京電力の電気料金を参考に, 1kWhあたり\$0.5とし, 測定した消費電力量を掛け計算した. これら3つのグラフをしてみると, 図4においては, level Iとlevel IIの処理時間が等しいにも関わらず, 図5ではlevel Iの方がlevel IIに比べ, 従量制料金が高くなっている. これは, 今回投入したジョブは, level IIまでの状態で十分に負荷分散可能なジョブであり, level Iでは, 無駄にパブリッククラウドのリソースを使ってしまうためだと考えられる. それ以上のレベルにおいては, レベルが上がるに連れ負荷分散されにくくなっていることが, 図6の従量制料金が低くなっている点から読み取れる. それに伴って, ジョブの処理時間もかかっていることがわかる. また, 図6は図4と形状が似ていることから, 消費電力量は処理時間が多大な影響を与えていることがわかる. このように, 本ミドルウェアで飽和判断レベルを変化させることで, ジョブの実行時間とパブリッククラウドの従量制料金, プライベートクラウドの消費電力量料金を制御できる.

4. ミドルウェアの評価

測定した結果をもとに本ミドルウェアの評価を行う. 本ミドルウェアでは評価指標として以下の式を考える.

$$Totalcost = F * T_t + (T_R * N_R * C_R + P_L * C_L)$$

T_t : Execution time of Total Jobs, T_R : Execution time of Public Cloud, N_R : Number of Instances used on Public Cloud, C_R : Charges of Public Cloud, P_L : Power Consumption on Private Cloud, C_L : Charges of Power Consumption on Private Cloud

係数 F は, ユーザが時間的コストと金銭的コストのバランスをどのように考えるかを定める値であり, 時間的コストを金銭的コストに換算するための係数である. 本評価で

は, 従量制料金と消費電力料金の価格設定による金銭的コストの議論を行う. そして, 上記式の定数 F の比率を変化させたトータルコストによる評価を行う.

4.1 金銭的コストに関する議論

本研究では, 金銭的コストをパブリッククラウドの従量制料金とプライベートクラウドの消費電力量料金で考えるが, この2つの金額は常に一定ではない. そこで, 本評価ではそれらの値を様々に変化させ考慮した結果, 電力単価: 従量制単価 = \$3.0:\$0.5, \$0.5:\$0.5, \$0.5:\$1.5 の場合を金銭的コストの代表例としてトータルコストの評価を行う.

4.2 トータルコストによる評価

トータルコストの評価では, 係数 F が重要となる. 今回の評価では, その係数を $1/20, 1/200, 1/2000, 1/20000$ と変化させた. つまり, $1/20$ では処理時間を最も重要視しており, $1/20000$ はパブリッククラウドの従量制料金とプライベートクラウドの消費電力量料金による金銭的コストを最も重要視する.

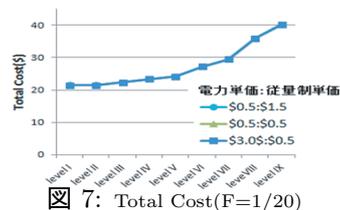


図 7: Total Cost (F=1/20)

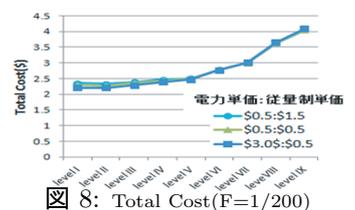


図 8: Total Cost (F=1/200)

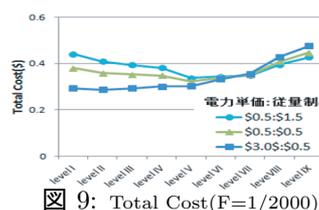


図 9: Total Cost (F=1/2000)

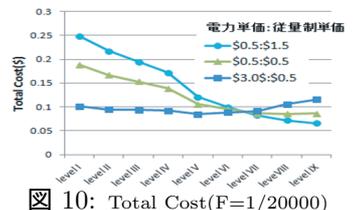


図 10: Total Cost (F=1/20000)

上記4つの図のように, それぞれの F の値を変化させる, つまりユーザが考えるコストを考慮した場合において, それぞれの結果でユーザの希望を満たすような level が存在することがわかる. このことから, 本ミドルウェアによって, ユーザが重要視するコストを低く抑えるような負荷分散を, パラメータを変化させることで実現できていることが確認された.

5. まとめと今後の課題

本研究では, ハイブリッドクラウド環境でデータインテンシブなジョブを効率よく処理しながら, さらにユーザのパラメータ設定により, 時間的コストや消費電力料金を含む金銭的コストを抑える負荷分散を行えるミドルウェアを提案し, 有用性を評価した. 今後は, その他様々なパラメータ設定での評価を行なっていきたい.

参考文献

- [1] Eucalyptus: <http://www.eucalyptus.com/>
- [2] 笠江優美子, 小口正人: 「ハイブリッドクラウド環境における省電力指向な負荷分散ミドルウェア開発に向けた一検討」, 第153回データベースシステム(DBS)研究会, B-3-2, 2011年11月