

# インタークラウドプロビジョニングに向けた計算リソースモデル

Towards Creating a New Computing Resource Model for Inter-cloud Provisioning

横山大作<sup>1</sup>

Daisaku YOKOYAMA

近山隆<sup>1</sup>

Takashi CHIKAYAMA

東京大学<sup>1</sup>

The University of Tokyo

## 1 はじめに

クラウドと呼ばれるオンデマンド型計算環境は、仮想化技術の進展とも相まって実用性が高まり、アプリケーション構築環境として近年急速に普及している。計算機を仮想的に必要なだけ利用できる IaaS (Infrastructure as a Service) と呼ばれるクラウド環境においては、Amazon EC2 をはじめとして国内外の多数のプロバイダがサービスを競う市場となっている。

IaaS は、仮想化された計算機そのものがユーザに提供されるため、利用する OS のレベルからユーザが決定できる自由度の高い計算環境である。複数のプロバイダが提供する様々な仮想計算機から、ユーザは自らのアプリケーション動作に望ましいものを自由に組み合わせて利用したいと考えている。例えば、通常時は安価なクラウドでアプリケーションを実行しているが、何らかの理由で突然アクセス数が増え、急激に負荷が高まった場合、通常時より高機能なクラウドを一時的に利用して性能を高め、突発的な需要に対応したい、という要求があり得る。また、利用していたクラウド環境に障害が発生した場合などには、現在利用可能な別のクラウドにアプリケーションを移動させ、サービスを継続させたいという高可用性に対する要求も存在する。

このように、クラウドを連携して利用したいという要求は高まっているが、実際に複数のクラウド環境を連携させて使うためには様々な課題が存在している。複数のクラウド環境を統一的に扱えるような仮想計算機管理システム、統一的なネットワークインフラ機構など、技術的に解決すべき課題に加えて、サービス提供アプリケーションの要求を満たせる計算環境構成をいかに発見し、もっとも望ましい構成をどのように選択するか、という点に解決しなければならない大きな課題が存在している。アプリケーションは通常、リクエストの 95% の応答時間を 1 秒以下にすること、などの何らかの目標サービスレベルを持っているが、複数の計算機環境が利用できる場合、現在利用しているものとは異なる計算機を利用したときに、どのような構成を取れば求めるサービスレベルを実現できるのか、どのような構成がもっとも低コストで要求を満たせるのか、を判断することは容易ではない。

我々は、このような問題に答えるため、複数クラウドを連携させることを想定したアプリケーション性能予測モデルを提案する。本モデルでは、あらかじめ取得したアプリケーション固有のスケラビリティ特性と、動作対象とする環境の持つ基礎的な性能特性とをもとにし

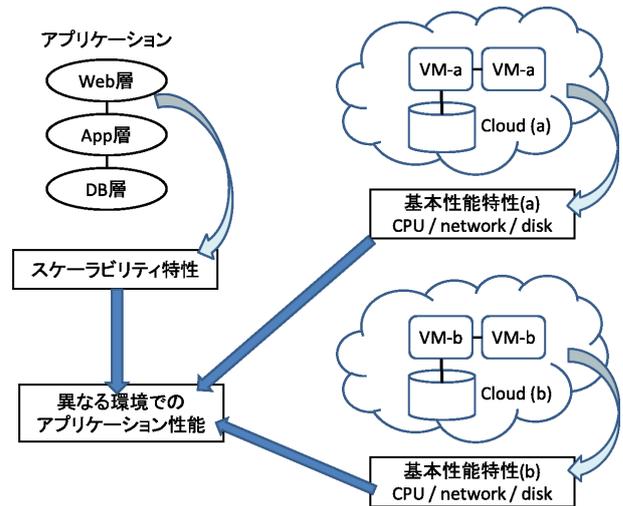


図 1 性能モデルの構築手順

て、動作環境でのアプリケーション性能を予測するという手順を取る。本モデルを用いて、既存のサービスアプリケーションを新しい実行環境上で動作させようと試みる際に、SLA を担保できる適切な構成が導出できることを目指す。ここでは、クラウドアプリケーションで広く用いられている multi-tier 型のアプリケーションを、この性能モデルの適用対象として考える。

本稿では、このような性能モデルの構築方法について述べ、複数クラウド環境を利用したリソースプロビジョニングへの利用への試みを紹介する。

## 2 性能モデル構築法

提案する性能モデルは、以下のような手順で性能を予測する。(図 1 参照)

1. 実行するアプリケーションを対象に、既知のクラウド環境において様々な機器構成を取った場合の性能変化を測定する。

対象アプリケーションは、クラウド上で動かすことが前提のものであるため、利用するリソース (機器) 構成を増加させることで性能を向上させることが可能になっている、スケールアウトを考慮した構成になっていると想定される。ここでは、多くのサービスアプリケーション作成時に利用される、multi-tier 構造のアプリケーションをモデル化の対象と想定する。multi-tier 構造の代表例として、web 層、アプリケーション層、DB 層の 3 層からアプリケーションを構成する 3 層モデルがよく知られており、広く用い

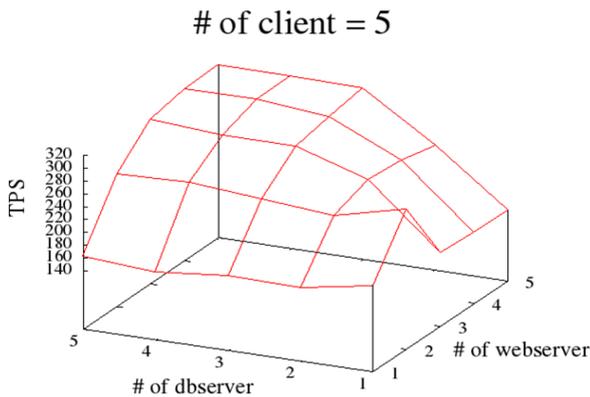


図2 アプリケーション性能測定例

られている。既知のクラウド環境において、それぞれの層を構成する仮想機械の数を変化させながらアプリケーションの性能測定を行い、アプリケーション性能の基本的なスケーラビリティ特性を把握することを試みる。

図2は、テスト用ウェブアプリケーションを対象に、ウェブサーバとDBサーバの数を変更したときの性能(トランザクション数/秒)変化を測定したものである。このような性能測定結果をもとに、ウェブサーバ数、DBサーバ数などを変数とするアプリケーションの相対性能モデルを作成し、機器構成を変化させたときの性能変化の度合いを把握することを試みる。

2. アプリケーションを実行する環境上でベンチマークソフトウェアを動作させ、環境の基礎的な性能値を得る。

複数のクラウド環境を利用しようとする際、アプリケーションはこれまで動作していた環境とは異なる性能の仮想機械上で動作することになる。この性能の違いを把握するために、ベンチマークソフトを用いて仮想機械環境の基礎的な性能情報を得ることを試みる。ベンチマークでは、

- CPU 性能
- ディスク I/O 性能
- ネットワーク I/O 性能

などを測定する。この性能測定は、(1)でアプリケーションのスケーラビリティ性能を測定した環境でも実施する。

3. (1)で求められたアプリケーションの性能変化傾向と、(2)で求められた種々の環境における基礎性能の差をもとに、求めたい環境での性能予測を行う。

以上のような方針でアプリケーションの性能をモデル化し、仮想機械構成の異なる複数のクラウドを利用する場合でも適切に性能予測が行えるような枠組みを構築することを目指している。

### 3 プロビジョニングへの利用

我々は、提案する性能モデルをリソースプロビジョニングに役立てることを目指している。すなわち、複数の

クラウド計算環境が利用可能であるとき、SLAを担保できるリソース利用計画にはどのようなものがあり、その中でどのリソース配置が望ましいのか、を見つけ出し、アプリケーションを実行しているユーザに提案することを目指している。

ある仮想機械構成によって実行されるアプリケーションがSLAを担保できるかどうかは、提案する性能モデルを用いて判断できる。現在利用可能なクラウド環境のリソースが入力されたとき、そこで実現可能な仮想機械構成群の中で、SLAを担保できる構成を探索すれば、性能面での制約を満たす構成を提示することが可能になる。利用可能な構成の自由度が増えるに従い、探索しなければならない空間が爆発的に広がることが考えられるため、探索空間の絞り込みを行うことによって現実的な時間内での構成案提示を実現することを目指している。

ここで、SLAを担保できる十分な性能の構成が多数得られることが想定されるため、何らかの基準によって「望ましい」構成から順に提示することが必要になる。望ましさを左右する大きな要素としては、クラウド環境利用時のランニングコストが考えられる。より安価にアプリケーションを実行できる構成が望ましいという考えは自然である。

しかし、ユーザはランニングコスト以外にも多様な価値基準を持っていることが想定される。「アプリケーションを管理する際に考慮することをなるべく減らすために、DB層だけは現在実行中の環境から動かしたくない」「アプリケーションの利用者がいる需要発生地とアプリケーションとの距離を、なるべく短く保ちたい」など、様々な価値基準が考えられるが、これらは必ずしも全ての局面でシステムが自動的に決定できるものではなく、状況に応じてアプリケーションを動かすユーザが方針を決定すべき事項である。

このような価値基準を反映するために、ユーザから、web層、DB層などのアプリケーション構成要素ごとの移動・変更のしやすさなどに関する評価基準を入力してもらい、ランニングコストと統合したコスト値を計算することで、構成の望ましさの基準としてプロビジョニングに利用することを試みている。

### 4 おわりに

我々は、異なる仮想計算機構成を利用する際のアプリケーション性能モデルを構築し、複数のクラウド環境を利用する際のリソースプロビジョニングに役立てることを目指している。今後、実用的なアプリケーションを利用した実証実験を通して、本手法の有用性の検証を行いたいと考えている。

### 謝辞

本研究の一部は、総務省委託研究「広域災害対応型クラウド基盤構築に向けた研究開発(高信頼クラウドサービス制御基盤技術)」において実施された。