

Repositório ISCTE-IUL

Deposited in *Repositório ISCTE-IUL*:

2018-06-08

Deposited version:

Post-print

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Belusso, C. L. M., Sawicki, S., Basto-Fernandes, V., Frantz, R. Z. & Roos-Frantz, F. (2017). Modelagem de preços de provedores de IaaS utilizando regressão múltipla. In Reis L.P., Rocha A., Alturas B., Costa C., Cota M.P. (Ed.), 12th Iberian Conference on Information Systems and Technologies, CISTI 2017. Lisbon: IEEE.

Further information on publisher's website:

10.23919/CISTI.2017.7975845

Publisher's copyright statement:

This is the peer reviewed version of the following article: Belusso, C. L. M., Sawicki, S., Basto-Fernandes, V., Frantz, R. Z. & Roos-Frantz, F. (2017). Modelagem de preços de provedores de IaaS utilizando regressão múltipla. In Reis L.P., Rocha A., Alturas B., Costa C., Cota M.P. (Ed.), 12th Iberian Conference on Information Systems and Technologies, CISTI 2017. Lisbon: IEEE., which has been published in final form at <https://dx.doi.org/10.23919/CISTI.2017.7975845>. This article may be used for non-commercial purposes in accordance with the Publisher's Terms and Conditions for self-archiving.

Use policy

Creative Commons CC BY 4.0

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in the Repository
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Modelagem de Preços de Provedores de IaaS Utilizando Regressão Múltipla

Price Modeling of IaaS Providers Using Multiple Regression

Cássio L. M. Belusso¹, Sandro Sawicki², Vitor Basto-Fernandes³, Rafael Z. Frantz², Fabricia Roos-Frantz²

¹Universidade Federal da Fronteira Sul, UFFS, Cerro Largo, Brasil

²Dartamento de Ciências Exatas e Engenharias, UNIJUI, Ijuí, Brasil

³Instituto Universitário de Lisboa (ISCTE-IUL), University Institute of Lisbon, ISTAR-IUL, Lisboa, Portugal
cassio.belusso@uffs.edu.br, sawicki@unijui.edu.br, vmbfs@iscte.pt, {rzfrantz, frfrantz}@unijui.edu.br

Resumo — Uma alternativa para usuários reduzirem custos de aquisição e manutenção de infraestrutura computacional para desenvolver, implementar e executar suas aplicações é a computação em nuvem. Os serviços de computação em nuvem são oferecidos por provedores e podem ser classificados em três modalidades: *Platform-as-a-Service* (PaaS), *Software-as-a-Service* (SaaS) e *Infrastructure-as-a-Service* (IaaS). Em IaaS, os provedores oferecem os serviços divididos em instâncias e o usuário tem à disposição uma máquina virtual com os recursos computacionais que desejar a um determinado valor. O principal desafio enfrentado pelas empresas é escolher, além do provedor, a instância que melhor se adapta as suas necessidades. Frequentemente, estas empresas precisam de uma grande infraestrutura computacional para gerir e aperfeiçoar seus processos de negócio e, diante do alto custo para manter uma infraestrutura local, têm migrado suas aplicações para a nuvem. Este trabalho busca fornecer subsídios capazes de auxiliar as empresas no processo de seleção do melhor provedor/instância para implantar e executar suas soluções de integração na nuvem. Para isso, um estudo preliminar para a elaboração de uma nova proposta de modelagem dos preços das instâncias de máquinas virtuais usando regressão linear é apresentado. Nesta abordagem são considerados os provedores Amazon EC2, Google Compute Engine e Microsoft Windows Azure.

Palavras Chave – computação em nuvem; modelagem matemática; modelo de preços; IaaS.

Abstract — An alternative for users to reduce costs of acquire and maintain computational infrastructure to develop, implement and execute software applications is cloud computing. Cloud computing services are offered by providers and can be classified into three modalities: *Platform-as-a-Service* (PaaS), *Software-as-a-Service* (SaaS) and *Infrastructure-as-a-Service* (IaaS). In IaaS, the providers offer services divided into instances. With this, the user has a virtual machine at their disposal with the computational resources desired at a given cost. The main challenge faced by companies is to choose what is the best pricing plan (instance/provider) to supply their computational demand. Frequently, these companies need a large computational infrastructure to manage and improve their business processes and, due to the high cost of maintaining local infrastructure, they have begun to migrate applications to the cloud. This work aims to provide insights that can help companies in selection process of

the best provider/instance to deploy and execute integrations solutions in the cloud. For this, a preliminary study to construction of a new proposal for price modeling of instances of virtual machines using linear regression is presented. In this approach, we consider the providers Amazon EC2, Google Compute Engine and Microsoft Windows Azure.

Keywords – cloud computing; mathematical modeling; pricing model; IaaS.

I. INTRODUÇÃO

A computação em nuvem é um dos principais avanços realizados no campo da tecnologia da informação atualmente. Neste novo conceito de computação, o usuário pode fazer uso das mais diversas aplicações através da Internet, como se elas estivessem instaladas em seu próprio computador.

Um dos principais atrativos da computação em nuvem é a relação custo-benefício, na qual o usuário paga somente pelo que utiliza [1]. Além disso, ao contratar um serviço na nuvem, o usuário já dispõe de serviços essenciais, como manutenção, atualizações, *backup* e segurança, os quais precisariam ser executados de forma manual, caso a aplicação estivesse instalada em qualquer tipo de dispositivo físico.

Os serviços de computação em nuvem são oferecidos por provedores e, geralmente, são resumidos em três modalidades principais: *Software-as-a-Service* (SaaS), onde o usuário utiliza determinado software e paga pela utilização; *Platform-as-a-Service* (PaaS), onde o usuário dispõe de um ambiente para projetar, testar e implantar aplicações personalizadas; e *Infrastructure-as-a-Service* (IaaS), onde o usuário contrata máquinas virtuais e gerencia os recursos computacionais como desejar. Esta pesquisa foca especificamente em IaaS.

Os provedores de IaaS oferecem um conjunto de diferentes planos para contratação denominados instâncias e que variam de acordo com as necessidades de cada cliente. Geralmente, pessoas físicas e até mesmo pequenas empresas necessitam de instâncias que contemplem pequenas demandas. As grandes empresas, por sua vez, possuem grandes demandas, e isso as tornam principais responsáveis pelos maiores investimentos em computação em nuvem.

Os preços das instâncias praticados pelos provedores são baseados na quantidade de recursos computacionais que as compõem. Porém, uma questão importante é compreender como estes preços são definidos. Algumas pesquisas demonstram que há, de fato, um mecanismo definido para isso [1,2,3]. Já outras mostram que a definição do preço é muito mais complexa e que sofre influência da política econômica e da lei da oferta e da procura [4,5]. Outros fatores ainda podem provocar alterações no preços das instâncias, como o local onde a máquina virtual está hospedada ou o sistema operacional escolhido pelo usuário. Diante disso, os usuários enfrentam um verdadeiro dilema durante o processo de tomada de decisão.

Em muitos casos, a tomada de decisão pode ocorrer de forma precipitada, pois o menor preço nem sempre pode ser a melhor escolha. Por exemplo, se um usuário contrata uma instância com mais recursos computacionais do que realmente necessita, ele estará gastando dinheiro desnecessariamente. Por outro lado, caso ele contrate uma instância mais barata buscando economia, o serviço pode ficar comprometido.

No contexto da Integração de Aplicações Empresariais (do inglês *Enterprise Application Integration* - EAI), a computação em nuvem proporciona uma infraestrutura computacional de alta capacidade a um baixo custo, nas quais as soluções de integração podem ser implantadas e executadas.

As soluções de integração são extremamente importantes no processo de integração, pois são softwares que atuam como um elo de comunicação entre as diferentes aplicações contidas no ecossistema de software das empresas, fazendo com que as mesmas possam compartilhar informações de forma rápida e eficiente [6]. A Fig. 1 ilustra o esquema básico de uma solução de integração.

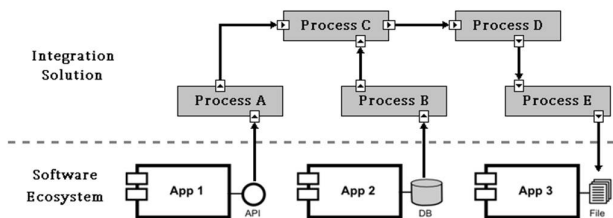


Figura 1. Solução de integração

Apesar das grandes empresas serem responsáveis por significativos investimentos na computação em nuvem, os provedores não disponibilizam um método que descreva, em detalhes, a variabilidade dos serviços e as restrições entre eles, de modo que estes modelos possam ser utilizados no processo de tomada de decisão [7].

Neste artigo, é apresentado um estudo preliminar para a elaboração de uma nova proposta de modelagem de preços das instâncias praticados por três provedores de computação em nuvem: Amazon EC2 (para o restante do texto somente Amazon) [8], Google Compute Engine (para o restante do texto somente Google) [9] e Microsoft Windows Azure (para o restante do texto somente Azure) [10]. Algumas hipóteses simplificadoras a serem adotadas neste novo modelo são discutidas e uma análise de preços de um pequeno grupo de instâncias do Azure é realizada.

O principal objetivo desta pesquisa é desenvolver um modelo capaz de auxiliar as empresas na otimização do processo de tomada de decisão durante a escolha do melhor provedor/instância para a migração de suas soluções de integração para a nuvem. Com isso, espera-se tornar a política de preços mais transparente para que as empresas possam reduzir o tempo e o custo da implantação e execução de suas soluções de integração em infraestruturas na nuvem.

O restante deste artigo está organizado da seguinte forma: a Seção II resume os trabalhos relacionados que apresentam propostas de modelagem de preços praticados por alguns provedores existentes no mercado, bem como algumas perspectivas de análise da política de preços praticada por eles; a Seção III apresenta um estudo de hipóteses para a elaboração de uma nova proposta e das teorias escolhidas para a construção do modelo; por fim, na Seção IV são apresentadas as conclusões e algumas perspectivas para trabalhos futuros.

II. TRABALHOS RELACIONADOS

Os trabalhos apresentados nesta seção buscam compreender a forma como os provedores de IaaS estabelecem a política de preços praticada para seus serviços. A percepção dentro da comunidade acadêmica é de que este entendimento não é tarefa fácil, o que pode justificar as diferentes abordagens adotadas nas pesquisas.

Murthy et al. (2012) apresentam um estudo comparativo dos modelos de cobrança e também do esquema de preços praticado por alguns provedores de IaaS. Os autores admitem como requisito de hardware principal o espaço de armazenamento, porém CPU e memória também são usados para comparar as instâncias. Eles também analisam diferenças nos preços das instâncias para dois sistemas operacionais diferentes e na forma de contratação do serviço. Por outro lado, a hospedagem das máquinas virtuais em diferentes localizações geográficas não são consideradas, mesmo este sendo um fator de grande influência nos preços.

Rodamilans et al. (2014) afirmam que o problema da seleção das instâncias pode ser resolvido por meio da caracterização das mesmas e também das aplicações, considerando o custo financeiro e o desempenho da máquina virtual. Os autores propõem uma metodologia para a seleção de provedores na nuvem por meio das seguintes etapas: Caracterização, Seleção e Execução. Para isso é admitida a hipótese de que não é preciso executar a aplicação em todas as instâncias para decidir a mais adequada, basta identificar o recurso computacional dominante.

Mazrekaj et al. (2016) destacam que, em uma oferta de serviços de computação em nuvem, o objetivo dos provedores é sempre obter a maior receita através de seus esquemas de preços, enquanto que os usuários buscam pela melhor QoS a um preço baixo. Para isso, eles fazem uma comparação dos preços de alguns modelos e esquemas de preços que são fornecidos pelos provedores de serviços em nuvem com base nos seguintes aspectos: serviços prestados, qualidade, preço justo e importância no mercado.

Chun e Choi (2014) analisam dois esquemas de preços praticados por provedores de computação em nuvem: *Subscription* e *Pay-per-use*. Eles afirmam que os diferentes

esquemas de preços praticados pelos provedores são consequência das diferentes estruturas de custo que os provedores possuem, custo este que parte desde o investimento na infraestrutura da nuvem até o custo de um serviço ficar ocioso por um determinado período de tempo.

Kihal et al. (2012) apresentam dois métodos para melhorar a transparência dos preços praticados pelos provedores de IaaS. O primeiro é o método de preços hedônico, em que é possível calcular a fração com que cada componente contribui no custo final da instância. Neste método, as instâncias precisam ser analisadas separadamente, impossibilitando a comparação entre provedores. O segundo método é o *Pricing Plan Comparison* (PriCo), que identifica o melhor perfil para cada provedor. Em ambos, os autores consideraram como hipótese principal que os clientes buscam o serviço mais barato. Durante a análise, foram utilizadas as componentes de IaaS como sendo memória, CPU e armazenamento, o Windows como sistema operacional e as instâncias *On-demand*.

Mitropoulou et al. (2016) propõem a construção de um índice de preços baseado no método hedônico de preços, que considera diferentes fatores presentes nos modelos de cobrança, dentre eles a localização geográfica do provedor. Trata-se de uma abordagem matemática, especificamente modelos de regressão (linear e exponencial, para efeito de comparação), onde o preço de um serviço está relacionado às suas características. Com o método hedônico, os autores tornaram possível o ajuste do preço de um serviço pela qualidade, não pela quantidade. Os autores coletaram dados de provedores e analisaram em quais regiões eles possuem máquinas virtuais hospedadas. O modelo também contempla as variáveis qualitativas referentes ao sistema operacional e ao modelo de contratação. As variáveis quantitativas referem-se à quantidade de memória, CPU, armazenamento e *transfer_out*. Segundo os autores, alguns fatores não foram considerados no modelo por não apresentarem influência significativa no preço.

A revisão do estado da arte mostra que não existem propostas diretamente voltadas à seleção de provedores de IaaS para a migração de soluções de integração para a nuvem. Da mesma forma, não foram encontrados instrumentos ou metodologias capazes de quantificar a demanda de recursos computacionais que uma solução de integração consome. Diante desta perspectiva, uma nova abordagem se faz necessária.

III. MODELAGEM PROPOSTA

Nesta seção, um conjunto de informações relevantes para a elaboração de uma nova proposta de modelagem de preços adotadas por provedores de IaaS é apresentado. O foco principal encontra-se voltado à EAI, especificamente na migração de soluções de integração para a nuvem. Devido a essa particularidade, uma análise de dados coletados e uma definição de hipóteses simplificadoras para o problema são discutidas nesta primeira etapa, considerando objetivos próprios e público-alvo.

O objetivo assumido na modelagem é obter sempre a instância com o menor preço, mas com uma Qualidade de Serviço (QoS) capaz de executar a demanda computacional pré-determinada.

A. Hipóteses Simplificadoras

Dentre os inúmeros provedores de IaaS existentes no mercado, neste trabalho optou-se pelo estudo de três deles, a saber: Amazon, Google e Azure. Algumas propostas são capazes de analisar mais do que um provedor ao mesmo tempo. Hernández et al. (2015) denominam esta capacidade de comparação *Cross-Provider*.

Diante da necessidade de hipóteses simplificadoras para uma nova proposta pode-se destacar, inicialmente, alguns fatores que influenciam nos preços praticados pelos provedores. São eles:

- Sistema Operacional: A opção por uma máquina virtual que utilize Linux ou Windows, por exemplo, pode resultar em uma diferença significativa no preço das instâncias. Em alguns provedores, o preço também pode variar se as plataformas forem diferentes. As informações coletadas diretamente do site dos provedores na Internet indicam que uma instância é sempre mais barata quando opta-se pelo Linux;
- Tipo de Instância: Existem instâncias compostas por configurações menores ou maiores, ou ainda instâncias do tipo *High-Memory* ou *High-CPU*, todas elas agrupadas de acordo com determinado recurso computacional. As instâncias *High-Memory* possuem grandes quantidades de memória, enquanto a quantidade dos demais componentes da instância é pequena ou tem pequenas variações. Da mesma forma, as instâncias *High-CPU* referem-se à grande capacidade de processamento. As soluções de integração são aplicações leves e não demandam de muito espaço de armazenamento. Neste sentido, esse requisito não é considerado. Alguns provedores possuem instâncias que diferem entre si quanto à unidade de armazenamento, que pode ser *Hard Disk Drive* (HDD) ou *Solid-State Drive* (SSD). Existem também o agrupamento de instâncias considerando os perfis dos usuários, onde os que possuem pouca utilização podem optar por instâncias com configurações menores. As instâncias com configurações maiores são indicadas para grandes empresas, pois as mesmas possuem um grande volume de aplicações executando simultaneamente;
- Localização Geográfica do Provedor: Muitos provedores possuem máquinas virtuais hospedadas em diversas partes do mundo e, dependendo da escolha, o preço da instância pode apresentar mudanças para o mesmo provedor. Existem provedores com muitas possibilidades de localização, dificultando a comparação. Nem sempre é possível optar pelo local onde a instância é mais barata, pois o usuário pode esbarrar em questões legais do país no qual a máquina virtual está hospedada. Nesta nova proposta, que tem como foco o menor preço somado à instância com a QoS ideal, uma pesquisa adjacente para cada um dos provedores deve ser feita com o intuito de identificar as regiões com os menores preços praticados, limitando-se a, no máximo, cinco delas;

- Modelo de Cobrança:** Além de possuírem inúmeras instâncias compostas por diferentes configurações, os provedores oferecem também modelos de cobrança diferenciados. Por exemplo, a Amazon oferece três modelos de cobrança: *On-demand*, *Reserved* e *Spot*. Nas instâncias *On-demand*, o usuário paga de acordo com a utilização, sem compromisso a longo prazo; nas instâncias *Reserved*, o usuário contrata a instância por um determinado tempo e, por causa disso, paga um valor menor na taxa temporal de utilização (que pode ser por hora ou por minuto); por fim, as instâncias *Spot* permitem que o usuário participe de uma espécie de leilão pela capacidade computacional não utilizada. No entanto, o preço *spot* oscila com base na oferta e na demanda por capacidade disponível. Caso a oferta do usuário esteja acima do preço *spot*, a instância estará executando; do contrário, o serviço é interrompido. A necessidade de destaque no mercado da computação em nuvem faz com que os provedores ofereçam diferentes modelos de cobrança, dentre os quais o modelo *On-demand* e o modelo *Reserved* são os mais comuns. Os modelos *Reserved* são sempre mais baratos no longo prazo e o preço tende a cair ainda mais à medida que o tempo de assinatura aumenta;
- Período de Utilização:** Por fim, um estudo mais detalhado busca identificar se, dentre os provedores escolhidos, todos alteram seus preços para a utilização dos serviços em horários diferenciados e se esta alteração é significativa. Pode ocorrer do usuário optar por executar suas aplicações em horários diferenciados para conseguir uma redução do custo da sua máquina virtual. Nos horários considerados de pico, o preço das instâncias tende a ser maior. A definição dos horários de pico é feita pelo próprio provedor.

B. Análise de Dados

Apesar das hipóteses simplificadoras abordarem os três provedores citados na proposta, uma análise de preços mais detalhada foi realizada apenas para o Azure, cuja escolha se deve apenas por ter sido o único até o momento a ter os preços estudados de forma mais detalhada.

Durante o processo de coleta de dados, todas as possibilidades de localização geográfica para hospedagem da máquina virtual foram consideradas, bem como as informações referentes aos sistemas operacionais Linux e Windows. O modelo de cobrança adotado foi o *On-demand*, pois ele é o mais atrativo para a maioria dos perfis de usuário. Neste modelo, não existe compromisso a longo prazo e o usuário pode cancelar o serviço a qualquer momento.

Diante da grande quantidade de dados obtidos, retirou-se uma pequena amostra composta por um grupo de cinco instâncias, as quais foram escolhidas considerando a disponibilidade em um número maior de localizações geográficas. Fez-se a simulação de contratação de cada uma delas optando pela região Leste dos EUA, devido a ela ser um dos locais mais baratos e também pela disponibilidade de um número maior de instâncias. O sistema operacional adotado foi o Windows, pois é o mais utilizado atualmente, e a camada de preços foi a Padrão, pois ela oferece uma maior flexibilidade.

As instâncias do Azure são formadas pelos seguintes recursos computacionais: CPU, memória e armazenamento. A cobrança é feita por hora de utilização. Os dados são referentes ao mês de agosto de 2016.

O principal objetivo desta amostra foi analisar o comportamento do preço das instâncias mediante variação dos principais recursos computacionais. Para isso, gráficos de dispersão foram utilizados para verificar a forma de distribuição dos dados. As marcações presentes nos gráficos representam as cinco instâncias do Azure escolhidas, na ordem da menor para a maior.

Na Fig. 2 são apresentados os gráficos de dispersão da quantidade de CPU e de memória *versus* a taxa temporal de utilização, computada em horas. Através dos gráficos é possível perceber um comportamento linear para ambos os recursos computacionais em praticamente todas as instâncias, exceto na menor. Além disso, especificamente no gráfico da memória, nota-se uma leve mudança no coeficiente angular da reta na segunda instância.

Com o intuito de permitir uma melhor visualização, a representação do gráfico do armazenamento *versus* a taxa de utilização é feita separadamente, na Fig. 3. Pode-se destacar, assim como nos gráficos dispostos da Fig. 2, que o comportamento linear também está presente para este recurso computacional. Novamente, uma leve mudança no coeficiente angular da reta é detectada, porém, neste caso, isso ocorre somente na terceira instância.

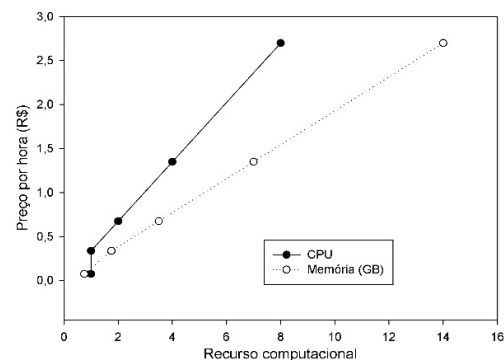


Figura 2. Gráfico de dispersão do recurso computacional *versus* preço cobrado por hora

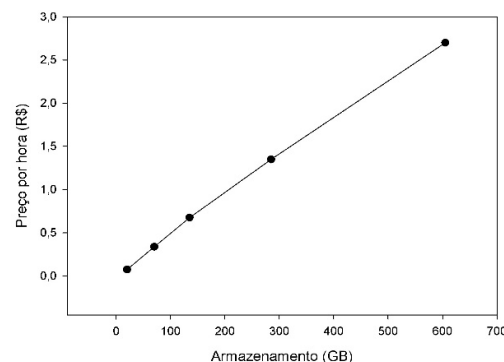


Figura 3. Gráfico de dispersão do armazenamento *versus* preço cobrado por hora

C. Modelo Matemático

Tendo em vista o comportamento linear dos dados coletados, optou-se pela regressão linear múltipla para estimar o preço das instâncias por haver mais de uma variável independente.

Através do método de regressão é possível estimar o custo individual das características que influenciam no preço final das instâncias por meio do cálculo dos coeficientes referentes a cada uma das variáveis do modelo. O modelo de regressão linear múltipla a ser adaptado para esta proposta encontra-se descrito pela Equação 1:

$$C_i = a_0 + a_1X_{i1} + a_2X_{i2} + \dots + a_nX_{in} + \varepsilon_i, \quad i = 1, \dots, n. \quad (1)$$

Onde:

- i refere-se à i -ésima instância, para qualquer um dos provedores escolhidos;
- n é o número de variáveis do modelo;
- C_i é o preço final da instância i ;
- a_i são os coeficientes da regressão a serem calculados, ou seja, a fatia de participação da característica X no preço final da instância i ;
- X_{in} são as variáveis independentes do modelo, no caso cada uma das características que influenciam no preço final da instância i ;
- ε_i é o erro residual da regressão para cada instância i .

As variáveis do modelo podem ser resumidas em variáveis quantitativas, compostas pelos requisitos de hardware (por exemplo, CPU e memória), e variáveis qualitativas, compostas pelos requisitos de software (por exemplo, sistema operacional) e pelas demais variáveis (por exemplo, localização geográfica). A representação das variáveis qualitativas do modelo são feitas por variáveis *Dummy*, em que uma determinada categoria pode assumir o valor 0 ou 1 [13]. A escolha das variáveis *Dummy* garante o comportamento linear da variável, motivo pelo qual não foram apresentados gráficos para as variáveis qualitativas.

Os coeficientes das variáveis quantitativas representam o preço unitário de cada um dos recursos, ou seja, o custo por cada unidade de memória, de CPU e de armazenamento.

D. Definição das Variáveis Dummy

As variáveis *Dummy* são definidas e inseridas previamente no modelo e representam as variáveis que não podem ser quantificadas. Nesta nova proposta de modelagem, por exemplo, é preciso definir a variável *Dummy* referente ao sistema operacional. Pode-se admitir que os valores 0 e 1 referem-se a duas categorias, Linux e Windows, respectivamente.

Para um melhor entendimento das variáveis *Dummy*, em uma situação onde o usuário opta pelo Windows como sistema operacional de sua máquina virtual, o coeficiente do modelo de regressão referente ao Windows é multiplicado por 1 (valor definido para Windows); se a opção for pelo Linux, o coeficiente é multiplicado por zero (valor definido para Linux).

Admitindo que uma instância é mais cara quando opta-se pelo Windows, o coeficiente desta categoria deverá ser positivo, gerando um acréscimo no preço final da instância; no caso da instância ser mais barata, o coeficiente deverá ser negativo, reduzindo o preço.

O mesmo raciocínio pode ser adotado para as demais variáveis qualitativas. No caso de uma variável qualitativa poder assumir mais de duas categorias (a localização geográfica da máquina virtual, por exemplo), o número de variáveis *Dummy* é $n-1$, sendo n o número de categorias. Para isso, é preciso estabelecer uma categoria de referência e definir as variáveis *Dummy* para as demais categorias.

IV. CONCLUSÕES

Neste trabalho foi apresentado um estudo preliminar para a elaboração de uma nova proposta de modelagem de preços das instâncias praticados por provedores de computação em nuvem a nível de IaaS utilizando um modelo de regressão linear múltipla. Trata-se de uma importante etapa no processo de criação de um mecanismo capaz de auxiliar as empresas no processo de tomada de decisão durante a migração de suas soluções de integração para uma infraestrutura na nuvem.

A revisão da literatura mostrou que nenhuma proposta encontra-se direcionada à EAI. A ausência de um mecanismo que possa oferecer às empresas a oportunidade de comparar os serviços oferecidos pelos provedores torna esta nova abordagem promissora, pois pode tornar a tomada de decisão uma tarefa mais simples e menos onerosa.

Algumas hipóteses para a criação desta nova metodologia foram discutidas considerando os provedores Amazon, Google e Azure. Além disso, gráficos de dispersão foram utilizados para justificar a escolha do modelo de regressão por meio de uma análise dos preços de um pequeno grupo de instâncias do Azure.

Para trabalhos futuros, espera-se realizar a análise de preços para os dois provedores restantes, Amazon e Google. Com isso, um modelo de regressão poderá ser implementado para cada um deles. Por se tratar de um método de estimativa, o nível de significância precisa ser considerado. É de senso comum que a maioria dos modelos apresentam não linearidades e, por causa disso, a modelagem pode ficar comprometida. Porém, com o iminente acréscimo de novas variáveis, o nível de significância tende a melhorar consideravelmente.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] M. K. M. Murthy, H. A. Sanjay, and J. P. Ashwini, "Pricing models and pricing schemes of IaaS providers: a comparison study," Proceedings of the International Conference on Advances in Computing, Communications and Informatics, pp. 143–147, 2012.
- [2] S. E. Kihal, C. Schlereth, and B. Skiera, "Price comparison for Infrastructure-as-a-Service," Proceedings of the European Conference on Information Systems, AISeL, 2012. <http://aisel.aisnet.org/ecis2012/161>.
- [3] P. Mitropoulou, E. Filiopoulou, C. Michalakelis, and M. Nikolaidou, "Pricing cloud IaaS services based on a hedonic price index," Computing, pp. 1–15, 2016.
- [4] M. Al-Roomi, S. Al-Ebrahim, S. Buqrais, and I. Ahmad, "Cloud computing pricing models: a survey," International Journal of Grid and Distributed Computing, vol. 6, pp. 93–106, 2013.

- [5] A. Mazrekaj, I. Shabani, and B. Sejdiu, "Pricing schemes in cloud computing: an overview," *International Journal of Advanced Computer Science and Application*, vol. 7, pp. 80–86, 2016.
- [6] R. Z. Frantz, R. Corchuelo, and F. Roos-Frantz, "On the design of a maintainable software development kit to implement integration solutions," *The Journal of Systems and Software*, vol. 111, pp. 89–104, 2016.
- [7] I. Hernández, S. Sawicki, F. Roos-Frantz, and R. Z. Frantz, "Cloud configuration modelling: a literature review from an application integration deployment perspective," *Procedia Computer Science*, vol. 64, pp. 977–983, 2015.
- [8] Amazon. 2016. "Amazon Web Services". [ONLINE] Available at: <https://aws.amazon.com/>. [Accessed 08 August 2016].
- [9] Google 2016. "Google Cloud Platform". [ONLINE] Available at: <https://cloud.google.com/>. [Accessed 06 August 2016].
- [10] Azure 2016. "Microsoft Windows Azure". [ONLINE] Available at: <https://azure.com/>. [Accessed 05 August 2016].
- [11] C. B. Rodamilans, A. Baruchi, and E. T. Midorikawa, "Experiences applying performance evaluation to select a cloud provider," *Proceedings of Recent Advances in Computer Engineering, Communications and Information Technology*, vol. 4, pp. 289–300, 2014.
- [12] S. H. Chun and B. S. Choi, "Service models and pricing schemes for cloud computing," *Cluster Computing*, vol. 17, pp. 529–535, 2014.
- [13] T. H. Wonnacott and R. J. Wonnacott, *Introductory Statistics for Business and Economics*, 4th ed., John Wiley & Sons: New York, 1990.