

Um Modelo de Custo para Nuvens IaaS baseado no Consumo de Energia de Máquinas Virtuais

Alternative Title: A Cost Model for IaaS Clouds based on Virtual Machine Energy Consumption

Mauro Hinz, Charles C. Miers, Maurício A. Pillon, Guilherme P. Koslovski
Programa de Pós-Graduação em Computação Aplicada (PPGCA)– UDESC
mmhinz@gmail.com, {charles.miers, mauricio.pillon, guilherme.koslovski}@udesc.br

RESUMO

A redução do consumo de energia de *data centers* é uma motivação constante para provedores de IaaS. Especificamente, o consumo de energia elétrica decorrente do uso de processadores é destaque entre os principais fatores consumidores. Entretanto, embora exista uma forte relação entre a carga computacional e o consumo dos processadores, os modelos de tarifação aplicados por provedores populares não consideram essa métrica como um elemento primário e variável. Esse trabalho quantifica esta relação, identificando o consumo por CPUs virtuais, que posteriormente são a base para rateio e cálculo do custo relacionado. O modelo proposto, Virtual Power, é discutido frente ao aplicado pela Amazon EC2. O estudo de caso aponta uma redução de custo para o provedor e um compartilhamento proporcional entre os usuários.

Palavras-Chave

Modelo, Custo, IaaS, Máquina Virtual.

ABSTRACT

Reduction in data center energy consumption is a constant motivation for IaaS providers. Among all components, CPU appears as a main energy consumer. Although there is a strong relationship between CPU load and its energy consumption, pricing models of popular IaaS providers do not consider this information as a primary and variable element. This paper quantifies the relationship by identifying the individual consumption of virtual CPUs, which form the basis for an allocation cost model. The proposed model, termed Virtual Power, is faced with Amazon EC2 pricing model pointing a cost reduction for IaaS provider and a proportional sharing between users.

Categories and Subject Descriptors

K.4.4 [Electronic Commerce]: Payment schemes; C.5.5 [Computer System Implementation]: Servers

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
SBSI 2016, May 17th-20th, 2016, Florianópolis, Santa Catarina, Brazil
Copyright SBC 2016.

General Terms

Management

Keywords

Cost, Model, IaaS, Virtual Machine.

1. INTRODUÇÃO

A demanda por serviços de computação está em constante aumento no decorrer das últimas décadas. Organizações e pessoas possuem necessidades crescentes por recursos computacionais para realizar as suas atividades profissionais e pessoais. Neste contexto, a forma como os recursos computacionais são ofertados e consumidos está em contínua evolução [14], sendo disponibilizados em locais remotos, como nuvens computacionais, que fornecem um modo automatizado de provisionamento aos seus usuários. Dentre os serviços ofertados por provedores de nuvem, IaaS (*Infrastructure as a Service*) disponibiliza máquinas virtuais (MV) sob demanda, configuradas de acordo com os requisitos dos usuários. IaaS é reconhecido como um dos serviços mais utilizados na migração de estruturas privadas para nuvens computacionais [17], motivado pelo fato de diversas empresas já possuem seus serviços encapsulados em MVs. Atualmente, provedores IaaS oferecem serviços usando um modelo de tarifação no qual o usuário paga somente pelo que utilizou. Esse modelo de tarifação não é inovador, já sendo empregado na oferta de recursos básicos como fornecimento de água e energia elétrica [18].

Os serviços de computação em nuvem são fornecidos por um conjunto de *data centers* projetados para compartilhar os recursos computacionais entre múltiplos usuários. Nesse cenário, o modelo ideal para tarifar a utilização dos recursos é uma questão em aberto. Estudos recentes investigam a relação de custos para manutenção do *data center*, sobretudo o consumo de energia elétrica [16]. Essa métrica ganhou importância por questões ambientais e aumento do valor de tarifação da energia elétrica [10].

Em *data centers*, o consumo de energia elétrica é tipicamente dividido em dois grupos: consumo geral e consumo relacionado à utilização dos equipamentos computacionais (e.g., servidores de processamento, dispositivos de armazenamento, redes) [1]. Especificamente, o consumo de energia elétrica decorrente do uso de processadores dos servidores figura entre um dos principais consumidores [5] [4] [14].

Observando a tarifação de uma MV aplicada por provedores de nuvens computacionais, pode-se constatar que o valor da tarifação não se altera conforme a carga de proces-

samento submetida. Isto é, se uma MV está usando 5% ou 100% do processador, o valor tarifado permanece constante. Sendo o processador de um computador um dos principais responsáveis pelo consumo de energia elétrica, é evidente a necessidade de um modelo de custo que considere a energia elétrica dependida para efetuar o processamento interno das MVs. Ainda, os modelos tarifários atuais não consideram um compartilhamento justo de custo entre os usuários, ou seja, uma MV que utilize menos recursos computacionais de um servidor deveria pagar um valor menor quando comparada a outra máquina virtual que consuma mais recursos.

Nesse contexto, este trabalho propõe Virtual Power, um modelo de custo baseado no consumo de energia do processador por MV em nuvens IaaS. Virtual Power é composto pelo consumo de energia individual de cada MV, somado ao custo compartilhado das operações realizadas pelos gerenciadores de MVs. A análise desta composição identifica a fração de energia consumida pela parte administrativa comum de um servidor hospedeiro, bem como o consumo de energia específico por uma MV. Em suma, as principais contribuições desse trabalho são: (i) o estudo da relação entre consumo elétrico de servidores e a utilização dos processadores físicos e virtuais; (ii) um modelo de tarifação guiado pelo consumo de energia do processador virtual; e (iii) a discussão sobre um estudo de caso comparando Virtual Power com o modelo de tarifação aplicado pelo provedor Amazon EC2.

O restante do trabalho está organizado da seguinte forma: a Seção 2 discute a relação entre o consumo elétrico e a utilização de processadores. O modelo de custo Virtual Power é apresentado na Seção 3 enquanto os trabalhos relacionados são discutidos na Seção 4. A Seção 5 descreve um estudo de caso aplicando o modelo Virtual Power. As considerações finais e perspectivas são apresentadas na Seção 6.

2. RELAÇÃO ENTRE O CONSUMO ELÉTRICO DE SERVIDORES E A UTILIZAÇÃO DOS PROCESSADORES

A demanda de processamento de um aplicativo executado em uma MV implica no mapeamento de uma CPU virtual (vCPU) desta máquina em uma CPU física do hospedeiro. Por sua vez, as instruções de execução deste aplicativo nesta vCPU implicam no processamento destas instruções na CPU física. Além das requisições oriundas das MVs, as tarefas básicas do *hypervisor* (e.g., gerenciamento de MVs e operações de entrada e saída) induzem em consumo de energia. Ainda, quanto maior o percentual de utilização do recurso virtual maior o consumo elétrico da máquina física [8].

Para aplicação do consumo elétrico na tarifação, é necessária a determinação exata do compartilhamento dessa métrica entre as MVs hospedadas por um servidor, considerando que as aplicações podem apresentar um consumo distinto de recursos de processamento e comunicação. Assim, linhas de base, necessárias para a elaboração do modelo de custo, são individualmente discutidas. A saber: consumos mínimo e máximo, relação entre o consumo de energia e a taxa de utilização do processador, e relação entre o consumo de energia e o volume de tráfego originado por MVs.

2.1 Ambiente de testes e metodologia

O ambiente de teste foi constituído por: (i) um equipamento identificado como SuperServer, com 2 processadores Intel Xeon E5-2600, 24 núcleos, 64GB RAM, 2TB SATA, fonte 740W redundante e 4 interfaces de rede In-

tel i350, executando o sistema operacional CentOS 64-bit, *kernel* 2.6.32.43-0.4.1, *hypervisor* XenServer 6.2 SP1. (ii) quatro hospedeiros HP, com processador AMD Phenom II X4 B93 2,8GHz, 4 núcleos, 4GB RAM, 500GB SATA, 2 interfaces de rede 8111E, fonte 150W, sistema operacional CentOS 64-bit, *kernel* 2.6.32.43-0.4.1, *hypervisor* XenServer 6.2 SP1. (iii) um nó coletor com processador AMD FX-6300 X6 3.5GHz, 6 núcleos, 16GB RAM, 1GB SATA, uma interface de rede Realtek 8111E, executando Debian Wheezy 7.0, *kernel* 3.2.34.27-0.1.1. (iv) *switch* D-link DGS-3100-24. (v) wattímetro para medição energética não-intrusivo com interface de leitura USB. O wattímetro foi posicionado entre a fonte externa de alimentação e o equipamento monitorado, contabilizando o consumo e repassando ao nó coletor.

Para emular cargas de ambientes IaaS, selecionamos: (a) *stress*¹: *benchmark* para processador e dispositivos de entrada e saída. (b) *iperf*²: teste de largura de banda em redes. (c) *cpulimit*³: ferramenta para limitar o uso do processador por processos. (d) *Peter Bailey watts-up*⁴: ferramenta para monitoramento e leitura dos dados do wattímetro. (e) *XenServer metrics*: ferramenta para monitoramento de recursos.

Para cada cenário experimental, os resultados das medições correspondem a 10 amostras, sendo cada ponto apresentado nos gráficos a média por um período de 7 minutos (com intervalo de 1 segundo). São desconsiderados os 2 primeiros minutos para eliminar leituras enquanto a carga de trabalho aplicada ao processador não está estabilizada. Quanto ao período total de medição, os cenários vinculados à taxa de utilização do processador tem periodicidade de 5% de crescimento pela carga de trabalho, enquanto aqueles vinculados a quantidade de dados trafegados por MV tem períodos iniciais entre 100 a 1000 Mbps e após de 1000 a 11000 Mbps.

2.2 Consumo de energia: mínimo e máximo

Os consumos mínimo e máximo de cada plataforma podem diferir, impactados pelo *hardware* e versão do *hypervisor*. A medição do primeiro ocorre quando o hospedeiro encontra-se sem MVs ativas, porém com os seus serviços básicos em execução, e do consumo máximo quando o *hypervisor* atinge 100% da potência de processamento do hospedeiro. Para o hospedeiro HP, foram identificados os valores 74 (mínimo) e 163 watts (máximo), enquanto o hospedeiro SuperServer informou 116 (mínimo) e 201 watts (máximo). Os consumos mínimos dos hospedeiros HP e SuperServer equivalem a aproximadamente 46% e 58% do consumo total, respectivamente. Observa-se que o consumo mínimo do SuperServer é 56% superior ao consumo mínimo do HP. Esta variação é justificada pela diferença arquitetural entre os hospedeiros (SuperServer possui arquitetura otimizada para servidor de processamento enquanto HP é um *desktop*).

O consumo máximo do SuperServer também foi superior ao do HP em 23%, atribuído ao número superior de núcleos de processamento, 24 contra 4 do HP. Porém, o hospedeiro SuperServer obteve melhores resultados se observada: (i) a relação entre consumo mínimo e consumo máximo (73% frente a 116% do HP); e (ii) relação consumo máximo de watt por núcleo com 8.3 contra 40 do HP. Pode-se concluir que mesmo com maior consumo de energia, tanto mínimo quanto máximo, o SuperServer tem melhores resultados nas

¹<http://people.seas.harvard.edu/~apw/stress/>

²<http://sourceforge.net/projects/iperf/>

³<http://cpulimit.sourceforge.net/>

⁴<https://www.wattsupmeters.com/>

relações entre consumo e potência de processamento. *Em suma, nos dois hospedeiros o consumo elétrico mínimo é representativo, ou seja, existe um custo coletivo que deve ser considerado em um modelo baseado em pay-as-you-use.*

2.3 Relação entre o consumo de energia e a taxa de utilização do processador

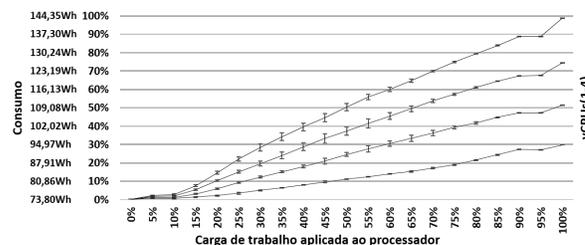
Os consumos de energia mínimo e máximo, apresentados na Seção 2.2, identificaram os limites inferiores e superiores do consumo de energia proveniente do uso do processador. Entretanto, para aplicação em um modelo de custo, uma relação gradual entre o consumo de energia e a taxa de utilização do processador é necessária. Os Cenários 1-1 e 1-3 mensuram o *hypervisor* dos hospedeiros HP e SuperServer, respectivamente; enquanto o Cenário 1-2 atua na medição em uma única MV com 4 vCPUs no hospedeiro HP; e o Cenário 1-4 mensura uma única MV com 24 vCPUs no hospedeiro SuperServer. Assim, os Cenários 1-1 e 1-3 representam o consumo de energia do *hypervisor* e os Cenários 1-2 e 1-4 estimam o consumo de energia por vCPU associado à MV.

Os resultados para os cenários envolvidos neste experimento são apresentados na Figura 1. Os gráficos têm no eixo X a carga de trabalho aplicada ao processador (entre 0% a 100%) e no eixo Y o valor do consumo de energia em watts, tendo como base o consumo mínimo (0%) e o máximo (100%). Na Figura 1(a), Cenário 1-1, o aumento gradual do consumo de energia segue o crescimento da carga de trabalho aplicada. A ascendência é mais perceptível a partir da carga de trabalho de 10%, com a qual o comportamento passa a ser linear até 100%, excetuando-se o ponto com carga de trabalho de 95%. A maior variação nas medidas ocorreu com a carga de trabalho de 45% e 3 vCPUs, apresentando um desvio padrão de 2,63%. Neste cenário, o consumo máximo no equipamento HP com carga de trabalho de 100% e 4 vCPU foi de 144 watts, ou seja, inferior ao consumo máximo total desta máquina que é de 163 watts.

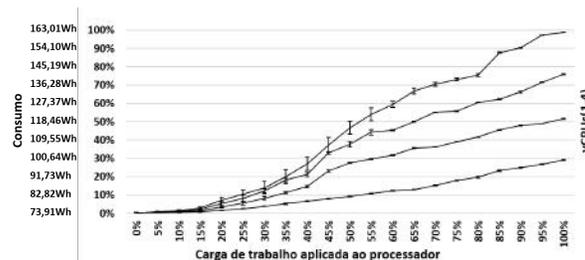
Os resultados referentes ao Cenário 1-2 são resumidos na Figura 1(b). Embora o gráfico tenha uma tendência linear, a ascendência gradual é menos estável quando comparada ao Cenário 1-1. A base de crescimento iniciou-se com carga 15% e apresentou alguns desvios entre os pontos com carga de trabalho 35% e 45%. Os resultados com 4 vCPUs apresentaram desvios padrão similares, sendo o maior de 4,10%. A maior carga de trabalho aplicada, de 100% com 4 vCPUs, atingiu o limite superior de consumo de energia, ou seja, o consumo máximo do hospedeiro HP de 163 watts.

Analisando os Cenários 1-1 e 1-2 é possível estabelecer a relação direta e gradual entre o consumo de energia e a taxa de utilização do processador. O impacto proveniente do uso do processador é desprezível para cargas inferiores a 15% e pode-se afirmar que *o consumo de energia está vinculado ao número de CPUs físicas ativas e não a vCPUs alocadas.*

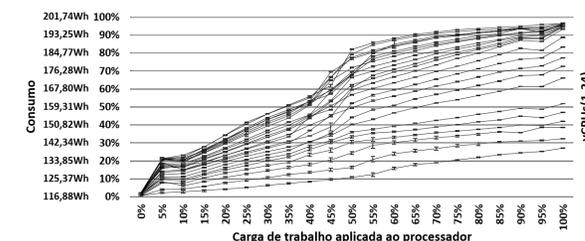
Os resultados obtidos nos Cenários 1-3 e 1-4 (hospedeiro SuperServer) estão representados nas Figuras 1(c) e 1(d). No Cenário 1-3, a carga de trabalho é aplicada diretamente no *hypervisor*, resultando em um aumento gradual do consumo de energia guiado pela carga de processamento. A ativação de um núcleo é representativa no consumo de energia: mesmo com a aplicação da menor carga de trabalho (5%) o resultado obtido para 24 núcleos ativos representou um incremento de aproximadamente 15% ao consumo mínimo. O comportamento possui tendência linear em duas faixas, a primeira com carga de trabalho entre 10% e 45% e



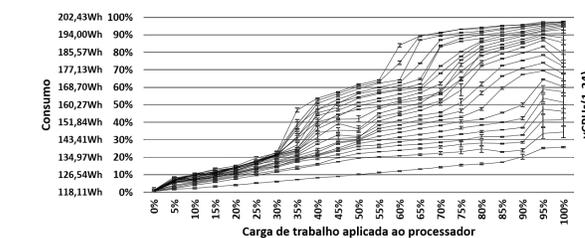
(a) Cenário 1-1.



(b) Cenário 1-2.



(c) Cenário 1-3.



(d) Cenário 1-4.

Figura 1: Relação entre o consumo de energia e a taxa de utilização do processador.

a segunda entre 55% e 100%, entretanto as faixas apresentaram taxa de crescimento distintas. Na primeira faixa, o incremento do consumo segue o incremento da carga, porém com pontos onde um menor número de vCPUs consumiram mais do que outros com um número maior de vCPUs. A segunda faixa tem taxa de crescimento menos expressiva pois aproxima-se da saturação de carga dos núcleos.

O Cenário 1-4 identifica a relação entre o consumo e a taxa de utilização do processador, tendo como hospedeiro o SuperServer e a aplicação das cargas de trabalho nas MVs. O comportamento (Figura 1(d)) segue os resultados anteriores, entretanto a carga de trabalho tem maior influência

no consumo de energia do que o número de núcleos ativos, se comparado com o Cenário 1-3. No Cenário 1-4, o incremento de 15% ao consumo mínimo só foi atingido com carga de trabalho de 25%, enquanto no Cenário 1-3 este incremento foi atingido com carga de trabalho de 5%. Um crescimento gradual menos acentuado pode ser percebido na Figura 1(d) para a faixa entre 0% e 30%. A partir da carga de 35% é possível observar um crescimento acentuado até o ponto de saturação dos núcleos entre 95% e 100%. O maior desvio padrão (5,35%) encontrado no Cenário 1-4 ocorreu com a carga de 100% e 12 vCPUS.

Os resultados apresentados nos Cenários 1-1 a 1-4 estabelecem uma relação direta entre o consumo de energia e a taxa de utilização de processador para dois hospedeiros distintos, tanto no hypervisor quanto nas MVs com um número de vCPUs distinto ativas. Desta forma, é possível estimar a quantidade de energia consumida observando-se a taxa de utilização do processador.

2.4 Relação entre o consumo de energia e a vazão de dados por MV

Em ambientes IaaS, as MVs alocadas possuem interfaces de redes virtuais similares as reais. Nas arquiteturas investigadas, virtualizadas com paravirtualização [2], o *hypervisor* é o único responsável pelo gerenciamento das solicitações de entrada e saída no hospedeiro, mesmo quando originadas nas MVs. Dessa forma, a transferência de dados impacta no uso do processador do *hypervisor*, independente do solicitante [20], pois o encaminhamento de pacotes entre as interfaces físicas e virtuais é efetuado pelo gerenciador. Assim, o consumo de energia para a transferência de dados é representado pelo consumo de energia do processador do *hypervisor* para a transferência de dados na rede em Mbps.

Para estabelecer a relação entre o consumo de energia e a quantidade de dados trafegados por uma interface virtual, 3 equipamentos HP atuaram como servidores *iperf*, enquanto os clientes, executados no hospedeiro SuperServer ou HP, geraram cargas constantes e crescentes entre 100 e 11000Mbps. Nos Cenários 2-1 e 2-2, o par de comunicação cliente-servidor encontra-se em MVs distintas, porém hospedadas na mesma máquina física, HP no Cenário 2-1, e SuperServer no Cenário 2-2. Nos Cenários 2-3 e 2-4 o par de comunicação encontra-se em máquinas físicas distintas: os servidores executam nos equipamentos HP (sem virtualização) e os clientes em MVs nos hospedeiros, (HP no Cenário 2-3 e SuperServer no Cenário 2-4). Assim, o experimento estabelece a relação de consumo nos hospedeiros, desconsiderando o consumo dos servidores sem virtualização (Cenários 2-3 e 2-4). Nos cenários 2-1 à 2-4, o *hypervisor* possui somente uma vCPU.

Na Figura 2, os gráficos têm no eixo X a carga de trabalho aplicada a interface de rede (100 a 11000 Mbps) e no eixo Y o valor do consumo de energia em watts. No Cenário 2-1 observa-se um comportamento crescente e gradual do consumo de energia conforme o incremento da vazão de dados. Neste cenário, o hospedeiro é o HP e todas as MVs, tanto clientes quanto servidores, estão hospedados no mesmo recurso físico. Três situações são destacadas: (i) um único par de comunicação; (ii) dois pares de comunicação em execução simultânea; e (iii) três pares de comunicação em execução simultânea. O consumo de energia mensurado para cargas entre 100 e 500 Mbps atinge, para as três situações apontadas, valores inferiores a 5 watts apresentando um leve crescimento, para as mesmas três situações, no intervalo com

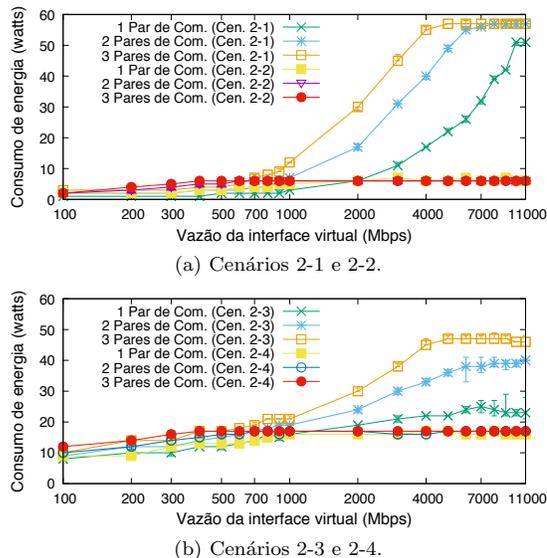


Figura 2: Relação entre o consumo de energia e a vazão de dados por MV.

cargas entre 600 e 1000 Mbps. O comportamento é alterado assim que a taxa de transferência ultrapassa o limite teórico da interface física (1000 Mbps) chegando a 60 watts, nas situações (ii) e (iii), e 50 watts na situação (i). Embora uma única máquina física hospede as MVs envolvidas, o encaminhamento de pacotes em uma arquitetura paravirtualizada requer a cópia de dados entre espaços temporários de armazenamento. Essas operações aumentam o processamento realizado pelo *hypervisor*, consequentemente aumentando seu consumo de energia. Quanto aos dados trafegados, a vazão agregada é menor que a capacidade total teórica devido à disputa no acesso as estruturas internas de encaminhamento [20]. O consumo de energia para o Cenário 2-2 é constante frente ao Cenário 2-1. Os servidores dos pares de comunicação estão em máquinas físicas exclusivas enquanto os clientes em MVs hospedadas no equipamento HP. Neste caso, todo o tráfego de dados entre os pares de comunicação passa pela interface de rede física do hospedeiro. O consumo de energia nas três situações (i), (ii) e (iii) com qualquer carga de trabalho não ultrapassa 7 watts, ou seja, 7,8% do consumo de energia do processador (máximo de 89 watts).

Na Figura 2(b) são apresentados os resultados com o hospedeiro SuperServer (Cenários 2-3 e 2-4). O comportamento do Cenário 2-3 é semelhante ao observado no Cenário 2-1, para as cargas de trabalho entre 100 e 1000 Mbps, o consumo de energia atingiu até 20 watts. Posteriormente, um crescimento acentuado ocorreu com cargas de trabalho entre 1000 e 4000 Mbps, e a estabilidade até 11000 Mbps. O máximo de consumo foi de 50 watts na situação (iii) com carga de trabalho 11000 Mbps. No Cenário 2-4 observou-se um consumo máximo de 18 watts. Ainda, no SuperServer a proporção entre o consumo de energia proveniente da interface de rede física e o consumo do processador é de 21%.

A análise dos resultados dos Cenários 2-1 a 2-4 indicam que o consumo de energia proveniente do tráfego de rede é diferente se o par de comunicação está ou não hospedado na mesma máquina física. Ainda, indica a existência de um aumento de consumo ocasionado pela utilização das interfaces

de rede virtuais.

2.5 Discussão e identificação de premissas

Baseado nas premissas de compartilhamento e isolamento de recursos em ambientes virtualizados [13], o consumo de energia elétrica de uma vCPU, mapeada em um processador físico, não gera consumo elétrico em outro processador físico. Todavia, atividades de gerenciamento (escalonamento e multiplexação) destes recursos são de responsabilidade do *hypervisor*, e o processamento destas atividades é independente do uso individual dos recursos nas MVs. Escalonar uma vCPU que terá taxa de utilização de processador de 100% ou de 0%, independente a que MV ela pertença, impacta na geração do mesmo consumo de energia no *hypervisor*. A taxa de utilização do processador nas MVs não impacta no uso do processador do *hypervisor*, porém o inverso ocorre com o gerenciamento dos recursos. O *hypervisor* gerencia os recursos físicos do hospedeiro associando-os e multiplexando-os a recursos virtuais vinculados à MVs. No caso de processadores, o *hypervisor* associa processadores físicos a vCPUs segundo seu critério de escalonamento. O processamento no *hypervisor* para escalonar 4 vCPUs a uma única MV ou a 4 MV distintas é o mesmo e, conseqüentemente, o consumo de energia correspondente também. O consumo de energia gerado pela taxa de utilização da vCPU associada a um processador depende única e exclusivamente do uso desta vCPU. Portanto, pode-se concluir que o consumo de energia é impactado pelo número de vCPUs ativas, independente do número de vCPUs associadas as MVs.

Em ambientes de virtualização é comum a existência de MV ativas porém sem carga de trabalho por um determinado período. O foco deste trabalho é a individualização do consumo de energia por MV o que torna essencial a reflexão sobre qual é o consumo de energia de uma MV ativa sem carga de trabalho. O gerenciamento de recursos no hospedeiro é feito sob demanda, ou seja, o *hypervisor* associa um processador a uma vCPU se e somente se a MV solicitar. Uma MV sem carga de trabalho não gera solicitações e, portanto, não deve ter vCPU associada. Como os experimentos apresentados demonstram que o consumo de energia é proveniente do uso dos recursos, MVs sem carga de trabalho não resultam em consumo extra de energia.

Ainda, seguindo a análise experimental, a Figura 3 apresenta os componentes do consumo elétrico em um ambiente virtualizado. (i) *consumo mínimo*: consumo elétrico do *hypervisor* quando não há MVs ativas; (ii) *consumo do gerenciamento das MVs*: consumo elétrico referente ao uso de recursos para operação realizadas pelo *hypervisor*, como por exemplo escalonamento, alocação, entre outros; (iii) *consumo de E/S*: consumo relacionado ao uso do processador pelo *hypervisor* para gerenciar operações de E/S das MVs; e (iv) *consumo das MVs*: consumo elétrico pelas MVs que utilizam o processador da máquina física.

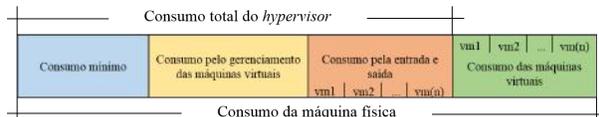


Figura 3: Componentes do consumo elétrico em ambientes virtualizados.

Por fim, a soma do consumo de energia elétrica gasto por duas vCPUs, com taxas de utilização de x e y , é equivalente ao consumo de energia elétrica de uma vCPU com taxa de utilização w , sendo $w = x + y$. Portanto, a proporção pode ser generalizada como indicado na Equação 1, na qual, m e n são o número de vCPUs e tx é a taxa de utilização.

$$m \equiv n \iff \sum_{i=1}^m tx_m = \sum_{i=1}^n tx_n \quad (1)$$

A análise experimental e as premissas teóricas identificadas motivam o desenvolvimento de um modelo de custo guiado pelo pela divisão proporcional do consumo elétrico de hospedeiros em nuvens IaaS. Neste sentido, o modelo deve definir como calcular cada componente envolvido (Figura 3) no provisionamento de uma MV em um serviço de nuvem IaaS, conforme discutido na próxima seção.

3. MODELO DE CUSTO VIRTUAL POWER

Virtual Power é um modelo de custo para recursos computacionais em nuvens baseado no consumo elétrico dos processadores alocados às MVs (a Tabela 1 resume a notação utilizada). O modelo tem como unidade de custo uma MV e seguindo a análise da Seção 2, pondera os componentes consumidores de energia (individuais e coletivos).

| Notação | Descrição |
|----------------|--|
| $C_r(u, i, j)$ | Consumo das operações de rede (interface j de uma MV i hospedada em u). |
| $C_{min}(u)$ | Consumo mínimo de um hospedeiro u . |
| $C_g(u)$ | Consumo de gerenciamento de u . |
| $vCPU_h(u)$ | vCPUs disponíveis em um hospedeiro u . |
| $vCPU_{vm}(i)$ | vCPUs solicitadas por uma MV i . |
| $C_t(u)$ | Consumo total de um hospedeiro u . |
| $U_{min}(i)$ | Custo mínimo proporcional atribuído a MV i . |
| $U_g(i)$ | Custo proporcional de gerenciamento atribuído a uma MV i . |
| $U_r(i)$ | Custo das operações de rede para MV i . |
| $U_p(i)$ | Custo da carga de processamento da MV i . |
| $U(i)$ | Custo total da MV i . |

Tabela 1: Notação para representar Virtual Power. As variáveis i e j foram utilizadas para indexar recursos virtuais e u para hospedeiros físicos.

O consumo de processamento de uma MV i e o consumo para suas operações de rede (C_r) são classificados como individuais porque eles contabilizam o consumo elétrico por MV, enquanto o consumo mínimo de um hospedeiro u (representado por $C_{min}(u)$) e o consumo de gerenciamento ($C_g(u)$) são considerados coletivos. Os consumos individuais fornecem os dados por MV, no entanto os consumos coletivos necessitam da definição de uma unidade de rateio. Considerando o fato de uso do processador ser o maior consumidor de energia do hospedeiro, a utilização de um processador virtual (vCPU) como a unidade de divisão dos custos de consumo coletivo é uma opção adequada, visando a justiça de compartilhamento dos recursos entre os usuários solicitantes. Embora outros critérios de rateio possam ser utilizados (e.g., número de MVs ativas no hospedeiro, vCPUS alocadas as MVs), Virtual Power buscou um compartilhamento de custo proporcional a utilização.

Considere um exemplo de alocação de vCPUs em processadores físicos, onde algumas MVs solicitaram a alocação de vCPUs, mas o *hypervisor* não as mapeou em processa-

dores físicos. Esse fato é comum no compartilhamento dos recursos em nuvens IaaS: um hospedeiro pode ter um número superior de vCPUs frente ao número de processadores físicos. Ainda, MVs podem ter vCPUs sem mapeamento real. O mapeamento de vCPUs é dinâmico, definido pelo escalonador do *hypervisor*.

Conforme discutido, o consumo mínimo de um hospedeiro é uma constante, obtida através de um método de calibração com equipamento de medição de energia no hospedeiro alvo. O rateio do consumo mínimo é descrito na Equação 2. $U_{min}(i)$ representa o consumo atribuído a uma MV i , enquanto $vCPU_h(u)$ e $vCPU_{vm}(i)$ representam o número de processadores virtuais disponíveis no hospedeiro u e o número de vCPUs da MV i , respectivamente.

$$U_{min}(i) = \left(\frac{C_{min}(u)}{vCPU_h(u)} \right) * vCPU_{vm}(i) \quad (2)$$

O consumo coletivo de energia elétrica necessário para gerenciamento das MVs é definido pelo consumo total do hospedeiro, decrescido do consumo mínimo e do consumo referente as operações de rede realizadas para as MVs hospedadas, conforme descrito na Equação 3. Na formulação, $C_t(u)$ representa o consumo de energia elétrica total do hospedeiro u enquanto $\#mv(u)$ e $\#int(i)$ representam o número total de MVs ativas no hospedeiro u e o total de interfaces de rede virtuais ativas na máquina virtual i , respectivamente.

$$C_g(u) = C_t(u) - C_{min}(u) - \left(\sum_{i=1}^{\#mv(u)} \sum_{j=1}^{\#int(i)} C_r(u, i, j) \right) \quad (3)$$

Seguindo a proposição de compartilhamento proporcional realizada para o consumo mínimo de um hospedeiro, o custo de gerenciamento é distribuído entre as MVs hospedadas (Equação 4). O custo de gerenciamento proporcionalmente atribuído a uma determinada MV i ($U_g(i)$) é calculado pela razão entre o custo de gerenciamento ($C_g(u)$) do hospedeiro u e o custo e o número de CPUs virtuais disponibilizadas pelo hospedeiro ($vCPU_h(u)$), multiplicado pelo número de vCPUs atribuídas à MV.

$$U_g(i) = \left(\frac{C_g(u)}{vCPU_h(u)} \right) * vCPU_{vm}(i) \quad (4)$$

O custo ocasionado pelas operações de rede realizadas por uma MV é obtido através do tráfego por interface de rede, conforme analisado na Seção 2. A Equação 5 descreve essa relação, quantificando o consumo das interfaces de rede de uma MV i hospedada por um recurso computacional u .

$$U_r(i) = \sum_{j=1}^{\#int(i)} C_r(u, i, j) \quad (5)$$

Definidas as políticas de rateio dos consumos coletivos e individuais, a Equação 6 identifica o custo individualizado de energia elétrica por MV, baseado na taxa de utilização dos processadores físicos.

$$U(i) = U_{min}(i) + U_g(i) + U_r(i) + U_p(i) \quad (6)$$

O consumo de energia por MV é constituído pelo somató-

rio dos consumos individualizados da máquina virtual, sendo $U_p(i)$ o custo resultante do processamento da MV. O resultado desta equação está em unidade de consumo de energia (e.g., watts). Assim, é necessário uma conversão para um custo com vínculo monetário, parametrizada pelo provedor de recursos. A precificação é uma opção do provedor IaaS, que pode atribuir valores distintos por faixas de horário. Essa discussão é atribuída a trabalhos futuros.

Virtual Power trata somente o componente energia da equação de custos totais de um provedor. O custo total pode crescer os custos com climatização, suporte, depreciação entre outros. Entretanto, esses custos são fixos e usualmente independentes da carga de processamento das MVs. Em suma, o modelo pode ser incorporado aos provedores públicos atuais (e.g., Amazon EC2, Microsoft, Google, Rackspace e Dualtec) que forneçam serviços IaaS. A Seção 5 apresenta um estudo de caso com o provedor Amazon EC2 exemplificando a aplicabilidade de Virtual Power.

4. TRABALHOS RELACIONADOS

A identificação dos componentes que compõem os custos, tanto fixos quanto variáveis, é um fator decisivo à correteza do modelo de custo. No ambiente de nuvem computacional, visto pela perspectiva do provedor, a cobrança dos serviços de alocação deve ser alimentada pelo custo dos componentes envolvidos na prestação de serviços. Dentre os trabalhos relacionados há pesquisas focadas na infraestrutura operacional e administrativa do provedor, bem como trabalhos focados no consumo elétrico pelo uso dos recursos.

Em [6] são apresentados três modelos de custo com foco na carga de trabalho aplicada ao processador. Cada modelo explora uma métrica distinta para inferir o custo relacionado com a carga de processamento. Uma proposta calcula o custo usando o consumo médio da carga de trabalho da aplicação, enquanto outra formulação calcula o custo utilizando a diferença entre o pico de utilização e carga média de trabalho da aplicação. Uma terceira abordagem calcula o custo utilizando como base os recursos não alocados. Contudo, o consumo elétrico não é abordado.

A relação entre o consumo de energia de processadores e discos de armazenamento quando submetidos a diferentes cargas de trabalho é apresentada no trabalho de [15]. Os autores indicam que a principal causa da ineficiência energética em *data centers* é o poder ocioso desperdiçado quando servidores executam em baixa utilização. Mesmo com uma carga de trabalho muito baixa (por exemplo, $\leq 10\%$ de utilização da CPU), a energia consumida é superior a 50% da potência de pico. Da mesma forma, se o disco, rede ou qualquer desses recursos é o gargalo de desempenho da aplicação, o desperdício de energia ociosa em outros recursos aumenta. O compartilhamento eficiente de recursos e a consolidação de servidores indicam vantagens introduzidas pela virtualização de recursos computacionais. Entretanto, a virtualização não eliminou a discrepância existente no consumo elétrico entre as MVs hospedadas, motivando a introdução do modelo Virtual Power. A relação individualizada do uso dos componentes de MVs e o consumo elétrico ocasionado foi alvo de estudo dos trabalhos [12] e [19]. O modelo de consumo de [8] determina o consumo elétrico pela utilização do processador, definindo a carga total de utilização pelo número de tarefas em execução. Este modelo não considera os componentes do ambiente virtualizado, sendo aplicável somente em ambientes com servidores físicos.

Considerando modelos para definição de custo em nuvens IaaS, alguns requisitos são identificados, vinculados ao consumo de energia dos recursos. São eles: (i) considerar recursos virtualizados. (ii) relacionar o consumo de recursos com o consumo elétrico. (iii) relacionar o consumo elétrico a tarifação. (iv) relacionar o consumo ao custo total de posse de um provedor, permitindo a identificação e avaliação de custos diretos e indiretos relacionados a disponibilização do serviço [9] [7]. (v) considerar o consumo ocasionado por operações de entrada e saída. A primeira característica específica se o modelo considera a virtualização de recursos computacionais, vinculando a carga ou uso do recurso virtual ao modelo, não considerada por modelos tradicionais [8] [6] [15].

Por sua vez, a relação entre o consumo de recursos e o consumo de energia é modelos como [8] e [15]. Estudos como [5] mostram que o consumo de energia é responsável por mais de 50% dos custos fixos de equipamentos de TI. Quanto aos itens (iii) e (iv), nenhum dos modelos analisados consideram estes requisitos na composição do custo, evidenciando uma carência na concepção de modelos adequados a necessidades de alocação de recursos sob demanda, sobretudo tarifados pelo seu consumo elétrico. Finalmente, embora os demais recursos de um equipamento de TI tenham menor impacto no consumo de energia, um modelo de custo que prime pela divisão proporcional dos custos vinculados ao consumo de energia deve considerar, pelo menos, o uso dos recursos com as operações de entrada e saída, especialmente tráfego de rede. O modelo proposto, Virtual Power, adere aos requisitos em sua totalidade.

5. ESTUDO DE CASO

Dentre os provedores IaaS públicos, a Amazon EC2 figura como uma infraestrutura popular. Para precificação dos recursos, a Amazon EC2 divulga uma tabela de custos por hora dos serviços fornecidos⁵. Entretanto, a composição do modelo de custo aplicado não é pública e, portanto, não é possível descrever detalhadamente seus componentes. Baseado nas informações disponíveis, essa seção apresenta um estudo de caso sobre o modelo de tarifação empregado pela Amazon EC2. O componente estudado é o custo de energia proveniente de equipamentos de TI que pode ser comparado com o modelo de Virtual Power. Amazon EC2 possui serviços de IaaS baseado no modelo de negócio *pay-as-you-go* e *pay-as-you-use*. A com *pay-as-you-go* define tarifas contratuais mensais fixas, independente se a MV está ativa ou não. O modelo de negócio *pay-as-you-use* precifica o serviço de IaaS alocado por hora e possui tarifação mensal variável, decorrente da ativação da MV, da quantidade de dados armazenados ou da quantidade de dados trafegados. Entretanto, é independente da taxa de processamento e consequente consumo de energia. Ambos modelos de negócio consideram implicitamente o custo com energia elétrica consumido pelo hospedeiro físico. Todavia, este custo não é detalhado e, principalmente, não é ponderado segundo o consumo efetivo do serviço locado. Assim, duas MVs do mesmo tipo, alocadas sob o modelo *pay-as-you-use* por uma hora, são tarifadas exatamente no mesmo valor, embora possuam carga de processamento distintas.

Como a informação sobre o valor do custo de energia por serviço não é disponibilizada, são considerados três percentuais de custo para comparação: 5%, 10% e 15%. Estes valores representam o custo total de energia por serviço

composta por sistemas de suporte e equipamentos de TI. Um levantamento realizado em vários *data centers* de médio porte nos EUA [5] indica que 52% do consumo de energia é oriundo dos equipamentos de TI e, destes, 15% são provenientes do uso de processador, base de cálculo do modelo de custo Virtual Power. Neste estudo de caso são comparados os valores de custos calculados pelo Virtual Power, em função da carga de trabalho da MV correspondente a 15% do custo total, com pontos percentuais fixos que representam algumas possíveis faixas de composição do custo de energia total do modelo de tarifação da Amazon EC2.

Dentre os tipos de MVs disponíveis na Amazon EC2, a configuração *m4.xlarge*⁶ foi selecionada por possuir características semelhantes ao hospedeiro SuperServer. Diante da impossibilidade de mensuração direta do consumo de energia do *m4.xlarge* e pela semelhança dos componentes físicos ao SuperServer, estabeleceu-se a premissa que o consumo mínimo e o consumo por núcleo do servidor no *data center* da Amazon EC2 é equivalente ao analisado na Seção 2. Considerou-se o valor de 116 watts como consumo mínimo e definiu-se a relação de consumo de energia por núcleo como proporcional com o intuito de se obter o consumo de energia segundo a carga de trabalho (0% à 100%).

O consumo total dos 24 núcleos do hospedeiro SuperServer é obtido pela subtração do consumo máximo (201 watts) pelo consumo mínimo (116 watts). O consumo por núcleo é o resultado da divisão do consumo total (85 watts) pelo número de núcleo (24), obtendo-se um consumo de 3,54 watts núcleo/hora. Considerando que o servidor *m4.xlarge* possui 16 núcleos e o mesmo consumo mínimo, tem-se um consumo máximo de 173 watts/hora. Na precificação disponibilizada pela Amazon EC2, existe uma variação de valores entre *data centers*. Esta variação ocorre devido a diferença do consumo de energia do *data center*, ocasionado pela localização geográfica ou pela infraestrutura, mas sobretudo, pela diferença no valor quilowatt-hora (kWh) da distribuidora de energia e impostos. Este estudo de caso baseou-se nos valores praticados para o *data center US West* localizado em Oregon, Estados Unidos. O valor do quilowatt-hora (kWh) comercial médio⁷ no estado americano é de US\$ 0,10.

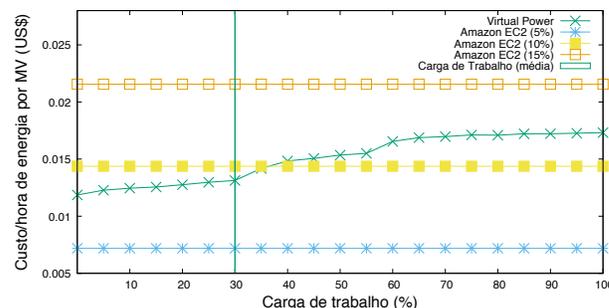


Figura 4: Custo de Energia por MV.

A comparação entre o custo de um serviço em IaaS hospedado na Amazon EC2 e o modelo de custo, baseado no consumo de energia do processador é resumido na Figura 4. No eixo X é indicada a carga de trabalho aplicada (0% a 100%) enquanto o eixo Y representa o valor cobrado por hora em dólar (USD). Três situações são constantes e representam o

⁵<https://aws.amazon.com/pt/ec2/pricing/>.

⁶<http://www.ec2instances.info/?filter=m4.4xlarge>.

⁷<http://www.statista.com/statistics/263492/>.

custo estimado atualmente tarifado na Amazon EC2, considerando o percentual de custo do consumo de energia em 5, 10 e 15. O comportamento variável representa a tarifação pelo modelo Virtual Power. Os trabalhos de [3] [11] indicam que a utilização típica dos servidores nos *data centers* está entre 10% a 50%, com média de 30%. Na Figura 4, se o percentual for de 15%, a Amazon EC2 tarifa o consumo de energia proveniente do processamento da MV acima do custo efetivamente gerado por este processamento, independente da carga de trabalho, e o oposto ocorre se o percentual for de 5%. Todavia, caso o percentual estipulado pela Amazon EC2 for 10%, o beneficiado dependerá da carga de trabalho aplicada. Para cargas de trabalhos inferior a 35%, o provedor estará tarifando mais do que efetivamente a MV estará gastando. O beneficiado se inverte a partir da carga de trabalho de 40%, onde o cliente passa a gastar mais energia do que está sendo tarifado. No entanto, a utilização típica dos servidores média é de 30%, ou seja, a Amazon EC2 tem a maior faixa de carga de trabalho como beneficiado, porém na maior parte dos casos, o cliente é beneficiado para o percentual de 10% também.

Neste estudo de caso, pode-se observar situações benéficas para provedores e clientes. O objetivo do modelo Virtual Power é permitir a tarifação proporcional para o modelo de negócio *pay-as-you-use*. Aplicar Virtual Power é economicamente viável para os clientes ou para o provedor. O ponto de equilíbrio neste estudo foi o percentual de 10 e carga de trabalho 35%, porém ao considerar que o modelo *pay-as-you-use* deve tarifar apenas pelo que é utilizado, Virtual Power é a escolha adequada em qualquer caso.

6. CONSIDERAÇÕES FINAIS

Os modelos de custo utilizados pelos principais provedores IaaS não consideram o percentual de uso de processadores, mas apenas a quantidade de processadores alocados por MV. Essa precificação não representa a carga atual de processamento e consumo dos recursos de um *data center*. Assim, uma precificação possível, quantificando o uso do processador e seu impacto no consumo de energia elétrica de um *data center* foi proposta no presente trabalho. O modelo proposto (Virtual Power) é guiado pela relação entre o uso de processamento das MVs e o consumo elétrico ocasionado. Virtual Power repassa proporcionalmente entre os usuários os custos com consumo elétrico de MVs. Em suma, o modelo é composto de custos individuais e coletivos, sendo o segundo proporcionalmente distribuído entre os usuários do recurso. A distribuição é guiada pelo uso real das vCPUs alocadas. Um estudo de caso comparou a aplicação de Virtual Power perante a precificação utilizada pela Amazon EC2. O estudo indicou que Virtual Power é capaz de diferenciar picos de uso e períodos de ociosidade, tarifando a reserva de acordo com a real utilização. Como trabalhos futuros, a aplicação de Virtual Power como métrica para balanceamento de carga e reconfiguração de *data centers* é uma linha promissora, bem como investigar e quantificar o impacto de outros recursos.

7. REFERÊNCIAS

- [1] V. Avelar, D. Azevedo, and A. French. Pue (tm): A comprehensive examination of the metric. Technical report, The Green Grid, 2012.
- [2] P. Barham, B. Dragovic, K. Fraser, S. Hand, T. Harris, A. Ho, R. Neugebauer, I. Pratt, and A. Warfield. Xen and the art of virtualization. *SIGOPS Oper. Syst. Rev.*, 37(5), Oct. 2003.
- [3] L. A. Barroso and U. Hözl. The case for energy-proportional computing. *Computer*, Dec. 2007.
- [4] V. Coroama and L. Hilty. Energy consumed vs. energy saved by ICT - a closer look. *23rd Int. Conf. on Informatics for Environmental Protection*, 2009.
- [5] ENP. Energy logic: Reducing data center energy consumption by creating savings that cascade across systems, 2008.
- [6] D. Gmach, J. Rolia, and L. Cherkasova. Resource and virtualization costs up in the cloud: Models and design choices. In *Proc. 41st IEEE/IFIP DSN*, 2011.
- [7] Y. Han. Cloud computing: Case studies and total cost of ownership. *Information technology & libraries*, 2011.
- [8] Y. Lee and A. Zomaya. Energy efficient utilization of resources in cloud computing systems. *The Journal of Supercomputing*, 60(2):268–280, 2012.
- [9] S. Li, K. Lim, P. Faraboschi, J. Chang, P. Ranganathan, and N. P. Jouppi. System-level integrated server architectures for scale-out datacenters. In *Proc. 44th IEEE/ACM MICRO*, 2011.
- [10] T. Mastelic, A. Oleksiak, H. Claussen, I. Brandic, J.-M. Pierson, and A. V. Vasilakos. Cloud computing: Survey on energy efficiency. *ACM Comput. Surv.*, 47(2):33:1–33:36, Dec. 2014.
- [11] D. Meisner, B. T. Gold, and T. F. Wenisch. Powernap: eliminating server idle power. In *ACM Sigplan Notices*, volume 44. ACM, 2009.
- [12] C. Mobius, W. Dargie, and A. Schill. Power consumption estimation models for processors, virtual machines, and servers. *Parallel and Distributed Systems, IEEE Transactions on*, June 2014.
- [13] G. J. Popek and R. P. Goldberg. Formal requirements for virtualizable third generation architectures. *Commun. ACM*, 17(7), jul 1974.
- [14] J. Shija, K. Bilas, S. A. Madani, and U. K. Samee. Data center energy efficient resource scheduling. *Springer Science+Business*, 2014.
- [15] S. Srikantaiah, A. Kansal, and F. Zhao. Energy aware consolidation for cloud computing. In *Proc. of the Conf. on Power Aware Computing and Systems*. USENIX, 2008.
- [16] M. K. Srinivasan, K. Sarukesi, P. Rodrigues, M. S. Manoj, and P. Revathy. State-of-the-art cloud computing security taxonomies: a classification of security challenges in the present cloud computing environment. In *Proc. ICACCI'12*. ACM, 2012.
- [17] W. Stallings. *Foundations of Modern Networking: SDN, NFV, QoE, IoT, and Cloud*. Addison-Wesley Professional, 1st edition, 2015.
- [18] S. Tang, B.-S. Lee, B. He, and H. Liu. Long-term resource fairness: towards economic fairness on pay-as-you-use computing systems. In *Proc. of the 28th ACM Int. Conf. on Supercomputing*, 2014.
- [19] D. Versick, I. Wassmann, and D. Tavangarian. Power consumption estimation of cpu and peripheral components in virtual machines. *SIGAPP Appl. Comput. Rev.*, 13(3), Sept. 2013.
- [20] P. Vicat-Blanc Primet, F. Anhalt, and G. Koslovski. Exploring the virtual infrastructure service concept in Grid'5000. In *20th ITC Specialist Seminar on Network Virtualization*, Hoi An, Vietnam, May 2009.