

# IoNCloud: uma abordagem não entrópica orientada a tráfego para reserva e isolamento de recursos em nuvens

Miguel C Neves<sup>1</sup>, Daniel S Marcon<sup>1</sup>, Rodrigo R Oliveira<sup>1</sup>,  
Leonardo R Bays<sup>1</sup>, Luciano P Gaspary<sup>1</sup>, Marinho P Barcellos<sup>1</sup>

<sup>1</sup>Universidade Federal do Rio Grande do Sul (UFRGS)  
Instituto de Informática

{mcneves, daniel.stefani, ruas.oliveira, lrbays, paschoal, marinho}@inf.ufrgs.br

**Resumo.** *Na computação em nuvem, a interferência entre tráfegos de naturezas diferentes gera imprevisibilidade no desempenho das aplicações. Propostas recentes, baseadas na formação de clusters virtuais isolando aplicações/locatários, resultam em subutilização de recursos ou aumento no custo de implementação. Neste artigo, propomos IoNCloud, um modelo de reserva e isolamento de recursos baseado nas características de tráfego das aplicações. Em IoNCloud, aplicações são atraídas ou repelidas de um mesmo cluster de forma a minimizar a interferência entre tráfegos. Experimentos demonstram que, em comparação com abordagens anteriores, IoNCloud acarreta baixa interferência entre tráfegos e reduz a subutilização dos recursos da nuvem. Ademais, IoNCloud não requer mudanças significativas na infraestrutura atual.*

**Abstract.** *The heterogeneity of traffic in intra-cloud networks leads to network performance variability (interference among traffic flows), which hinders application performance. Recent work addresses this issue by isolating applications/tenants in virtual networks, resulting in underutilization of resources or increased implementation costs. In this paper, we introduce IoNCloud, a traffic-based resource reservation and isolation model. In IoNCloud, an application is attracted to or repelled from a cluster according to the traffic interference it causes in other applications. Experiments show that, in comparison to previous work, IoNCloud leads to lower network performance interference and reduces the underutilization of resources. Moreover, IoNCloud does not require significant changes to current cloud datacenter infrastructures.*

## 1. Introdução

Utilização sob demanda e elástica de recursos e pagamento apenas pelo uso são alguns dos fatores que tornaram a computação em nuvem um paradigma atraente e em ampla difusão. Nesse contexto, benefícios como a redução dos custos com operação e manutenção da infraestrutura computacional fazem com que novos locatários migrem suas aplicações para os grandes *data centers* de nuvem [Armbrust et al. 2010]. Com o intuito de reduzir custos operacionais e viabilizar o modelo comercial proposto pelo paradigma, provedores de nuvem costumam empregar técnicas de multiplexação de recursos. Tal prática leva ao compartilhamento de recursos físicos por um elevado número de aplicações provenientes de diferentes locatários [Genez et al. 2012].

Recentemente, o desempenho baixo e imprevisível das aplicações, causado principalmente pela multiplexação dos recursos da rede interna da nuvem, tornou-se uma crescente preocupação [Ballani et al. 2011]. O fenômeno ocorre quando aplicações com tráfegos de naturezas diferentes compartilham o mesmo conjunto de dispositivos físicos [Benson et al. 2010]. Como alternativa, a literatura recente propõe o uso de virtualização de redes para isolar os locatários, de modo a oferecer garantia de recursos às aplicações e isolamento entre tráfegos [Guo et al. 2010, Ballani et al. 2011, Xie et al. 2012]. Tais estratégias levam à instanciação de uma rede virtual (ou *cluster* virtual) por aplicação ou por locatário. Embora reduza interferências, esse tipo de abordagem pode acarretar subutilização ou sobrecarga de gerenciamento dos recursos de rede da nuvem. Ademais, essas limitações são agravadas pela natureza variável dos requisitos de rede da maioria das aplicações hospedadas nos *data centers* de nuvem [Benson et al. 2010].

Este artigo propõe IoNCloud (*Isolation of Networks in the Cloud*), um modelo de reserva e isolamento de recursos de rede, como solução eficiente às questões de interferência entre tráfegos e imprevisibilidade de desempenho das aplicações. Em primeiro lugar, IoNCloud analisa os padrões de tráfego gerados pelas máquinas virtuais (*Virtual Machines* – VMs) das aplicações da nuvem, de forma a quantificar a necessidade de isolamento dos recursos de rede entre pares de aplicações. Em seguida, IoNCloud agrupa instâncias de aplicações e suas respectivas VMs em *Clusters* Virtuais Orientados a Tráfego (*Traffic-Oriented Virtual Clusters* – TOVCs), de acordo com as necessidades de isolamento. Esse procedimento leva à criação de grupos de aplicações de tal forma que a interferência entre tráfegos alocados a um mesmo TOVC seja a menor possível.

As principais contribuições deste artigo são resumidas a seguir:

- **Criação de métrica para expressar a necessidade de isolamento (Seção 3).** Definimos *entropia* como a variação em relação à média da demanda de determinado tráfego. Sendo assim, a métrica proposta quantifica quais aplicações apresentam melhor comportamento médio ao serem agrupadas sobre os mesmos recursos de rede.
- **Definição de um novo modelo para isolar tráfegos de rede em *data centers* de nuvem (Seção 4).** Amparado na métrica de entropia, IoNCloud permite melhor eficiência na utilização de recursos e redução na interferência entre os tráfegos das aplicações da nuvem. Além disso, o modelo independe da topologia de rede, de algoritmos de posicionamento de VMs ou de tipos de tráfego, tornando-se ortogonal à implementação de novas tecnologias nesse contexto.
- **Análise quantitativa do modelo proposto (Seção 5).** A eficiência de IoNCloud é analisada em cenários com base em um estudo de medição sobre tráfego em *data centers* [Benson et al. 2010]. Ademais, a abordagem é comparada com propostas anteriores, as quais não diferenciam aplicações ou tipos de tráfego.

O restante do artigo está organizado da seguinte forma. A Seção 2 apresenta os principais aspectos relacionados às características de tráfegos em *data centers* de nuvem. A Seção 3 define uma métrica para calcular a *entropia* da rede. A Seção 4 descreve o modelo de formação de TOVCs e agrupamento de aplicações, enquanto a Seção 5 faz uma análise deste modelo e compara-o com propostas da literatura. Os principais trabalhos relacionados são discutidos na Seção 6 e as considerações finais, apresentadas na Seção 7.

## 2. Tráfego e Causas de Interferência na Rede da Nuvem

Diversos fluxos de naturezas diferentes coabitam uma rede de *data center* de nuvem. Em geral, esses fluxos seguem uma classificação bimodal, onde a primeira classe representa aplicações com grandes volumes de dados, enquanto a segunda representa aplicações orientadas à interação com o usuário [Kandula et al. 2009]. Entretanto, a divisão dos fluxos em apenas duas classes é simplista, visto que dentro de cada grande classe existe uma variação significativa quanto ao comportamento dos fluxos [Benson et al. 2010].

Com o uso de mecanismos de multiplexação de recursos, aplicações com diferentes requisitos compartilham os dispositivos de rede. Em consequência, a rede interna da nuvem opera de maneira imprevisível, uma vez que tráfegos de naturezas distintas podem interferir uns nos outros [Abts and Felderman 2012]. Recentemente, a literatura vem identificando problemas particulares nos *data centers* de nuvem. Especificamente, tais problemas estão relacionados às características topológicas das redes de grande porte utilizadas nessas infraestruturas [Ballani et al. 2011].

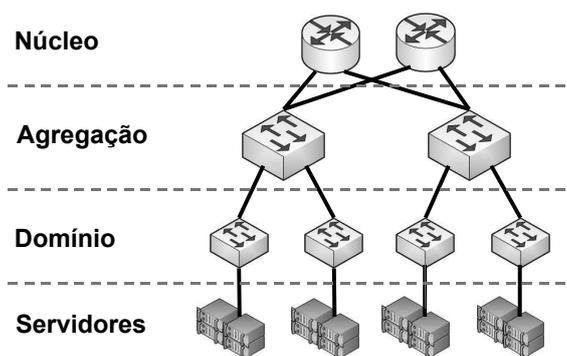


Figura 1. Topologia canônica de uma rede de *data center* de nuvem, adaptada de [Benson et al. 2010].

Conforme demonstrado na Figura 1, a topologia dessas redes está organizada em forma de árvore, com os servidores dispostos nas folhas e os demais dispositivos organizados de forma hierárquica. Como os dispositivos estão fisicamente muito próximos, o atraso de propagação é muito pequeno. Com isso, a maioria do tráfego em um *data center* pode ser caracterizado por um modelo *on/off* cujos períodos são de curta duração. Relacionado a isso, medições em [Benson et al. 2010] apontam que a utilização da maioria dos enlaces é pequena em todos os níveis da árvore exceto na raiz.

No entanto, apesar da baixa utilização, quando uma aplicação envia um fluxo de duração relativamente maior, o número de RTTs na conexão é suficiente para expandir a janela de transmissão do TCP, ao contrário das conexões de vida curta. Logo, o restante do tráfego nesse enlace acaba sendo prejudicado ao competir com a conexão mais longa. Mais precisamente, os fluxos mais curtos acabam sendo terminados antecipadamente ou geram retransmissões em demasia, agravando a situação [Abts and Felderman 2012]. Em última instância, atrasos e interrupções na propagação dos dados acabam por intervir na computação das VMs, prejudicando o desempenho geral das aplicações [Ballani et al. 2011].

A imprevisibilidade do tráfego das aplicações acarreta o uso ineficiente dos recursos das redes de *data centers* de nuvem. Dessa forma, é útil a definição de uma métrica que expresse a real necessidade de isolamento dos recursos de rede das aplicações hospedadas nesses ambientes. A próxima seção propõe uma métrica que expressa essa necessidade por meio da quantificação da interferência entre tráfegos.

### 3. Quantificando a Entropia na Rede da Nuvem

Neste artigo, denominamos de *entropia* ( $\mathcal{E}$ ) a variação em relação à média da demanda de determinado tráfego em uma rede de *data center* de nuvem. O objetivo desta seção é definir uma métrica, entropia, que quantifique a influência gerada por diversas instâncias de tráfego sobre a previsibilidade do tráfego agregado. Tais tráfegos são originados de aplicações distintas que, por uma circunstância do estado atual do posicionamento de VMs no *data center*, compartilham e competem por recursos de uma parte da infraestrutura. Considera-se que a interferência entre tráfegos é dada pela largura de banda excedente em relação à média do tráfego agregado. Sendo assim, uma contribuição chave deste artigo é a análise da entropia sobre o tráfego agregado, de forma a expressar com precisão a necessidade de isolamento das aplicações hospedadas na nuvem.

Conceitualmente, verifica-se qual o comportamento médio da demanda de tráfego resultante, ao serem agrupadas diversas aplicações sobre os mesmos recursos de rede. O cálculo do valor de  $\mathcal{E}$  é composto por três etapas: geração do tráfego agregado com base no histórico de demandas; discretização do tempo em  $N$  amostras para obtenção do desvio médio do tráfego agregado; e realização de procedimento de convolução, de forma a capturar variações temporais. A formalização de cada etapa é descrita a seguir.

Supondo  $\Pi$  como o conjunto de todos os tráfegos compartilhando os mesmos recursos de rede no data center, onde  $\pi \in \Pi$  representa o tráfego entre um par de VMs quaisquer. Sendo assim, o histórico da demanda por largura de banda de um tráfego qualquer  $\pi \in \Pi$  em determinado intervalo  $\Delta\tau = [t_i, t_f]$  é dado pela função  $h_\pi(t) : t \in \Delta\tau$ . Com isso, a demanda agregada  $\mathcal{H}(t)$  para um grupo de tráfegos compartilhando determinado enlace é dada pela soma de toda a demanda ao longo do tempo  $\mathcal{H}(\cdot) = \sum_{\pi \in \Pi} h_\pi(\cdot)$ .

Cada histórico de demanda  $h_\pi(\cdot)$  expressa a largura de banda consumida em um instante  $t$  ao longo do intervalo  $\Delta\tau$ . Pragmaticamente, isso implica a coleta de  $N$  amostras com intervalo de amostragem  $\delta t_A$  constante e pequeno o suficiente para expressar a precisão desejada<sup>1</sup>. Essa discretização do espaço amostral resulta em  $|\Pi|$  vetores na forma  $h_\pi[\cdot] = \{h_\pi[1], h_\pi[2], \dots, h_\pi[N]\}$ . Logo, o vetor  $H[\cdot] = \sum_{\pi \in \Pi} h_\pi[\cdot]$  expressa a demanda agregada por largura de banda ao longo do intervalo de tempo discretizado  $\Delta\tau$ . A seguir, são calculados a média  $M$  e o desvio médio  $D = \sum_{i=1}^N |H[i] - M|/N$  das amostras, expressando a variação do valor (agregado) ao longo do tempo.

Além de calcular a variação na demanda agregada, é preciso considerar que muitas vezes aplicações na nuvem não possuem um comportamento constante. Contudo, cada demanda de forma individual segue uma distribuição probabilística bem definida, indicando que os padrões de tráfego repetem-se com certa frequência [Benson et al. 2010]. Sendo assim, utiliza-se um procedimento de convolução com o intuito de agregar à métrica possíveis variações temporais no comportamento da demanda.

<sup>1</sup>Esse histórico de amostras pode ser obtido por meio de ferramentas de monitoramento como, por exemplo, o Amazon CloudWatch: <http://aws.amazon.com/cloudwatch>

O procedimento de convolução consiste em mover um dos tráfegos sobre todos os outros, de forma a capturar comportamentos deslocados ao longo do tempo. Para isso, inicialmente, todos os tráfegos são concatenados para gerar uma situação onde nenhum deles se intercepta. Essa concatenação resulta em uma janela de tempo de tamanho  $\Delta t_W$ . A seguir, para cada  $K = |\Pi| - 1$  iterações do procedimento, seleciona-se o último tráfego e desloca-o ao longo da janela de tempo  $\Delta t_W$ . A cada etapa de deslocamento, o cálculo do desvio médio é feito e a demanda do tráfego selecionado é adiantada em  $\delta t_C$  unidades, resultando  $\Delta t_W / \delta t_C$  operações. Ao final de cada etapa, o tráfego selecionado é descartado e  $\Delta t_W$  é atualizado. Com isso, seja  $\Delta t_{W_i}$  o tamanho da janela na iteração  $i \in K$  e  $D_{i,j}$  o desvio médio calculado na etapa de deslocamento  $j$  da iteração  $i \in K$ , a entropia é dada pela *média dos desvios médios* ao longo das  $K$  iterações do procedimento de convolução, conforme a seguinte equação:

$$\mathcal{E} = \frac{\sum_{i=1}^K \sum_{j=1}^{\Delta t_{W_i} / \delta t_C} D_{i,j} / (\Delta t_{W_i} / \delta t_C)}{K}.$$

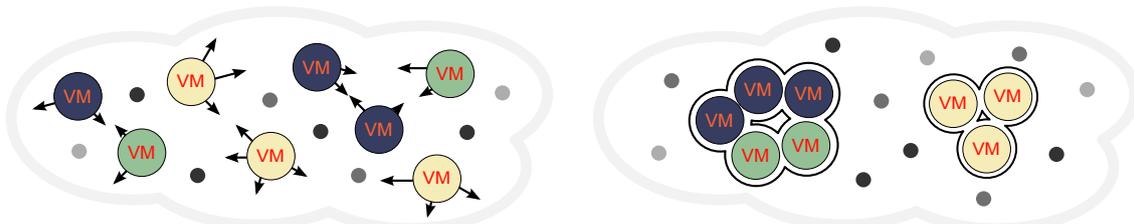
Essa métrica pode ser utilizada pelos sistemas de alocação e gerenciamento dos recursos da nuvem de forma a otimizar a reserva e o isolamento dos mesmos. Por esse motivo, ela forma o núcleo do modelo de alocação de recursos em *data centers* de nuvem proposto neste artigo, o qual será definido a seguir.

#### 4. IoNCloud

Conforme discutido anteriormente, o modelo proposto neste artigo visa prover isolamento de tráfego às aplicações de *data center* sem sacrificar a multiplexação de recursos da nuvem. Para alcançar esse objetivo, IoNCloud (*Isolation of Networks in the Cloud*) agrupa aplicações cujos tráfegos possuem pouca ou nenhuma interferência entre si, reduzindo a imprevisibilidade da nuvem. IoNCloud inspira-se em processos naturais (Figura 2), onde VMs (átomos) com alta interferência entre tráfegos aumentam a imprevisibilidade do ambiente, enquanto VMs com baixa interferência entre tráfegos são consequência de um ambiente estável, caracterizado por uma maior previsibilidade. Durante esse processo [Figura 2(a)], aplicações com grande entropia, ou seja, cujos tráfegos interfeririam um no outro, são repelidas, enquanto as demais são atraídas. O sistema torna-se estável [Figura 2(b)] quando a menor entropia global é atingida.

Mais especificamente, IoNCloud agrupa aplicações com baixa interferência entre si em um mesmo *cluster virtual orientado a tráfego* (definido a seguir, na Seção 4.1). A abordagem descrita neste trabalho (Seção 4.2) é complementar ao processo de alocação de recursos na infraestrutura da nuvem. Esse último visa instanciar VMs e/ou topologias de rede virtuais em recursos físicos, assumindo (i) que todas as aplicações devem ser alocadas no mesmo substrato; ou (ii) que estas devem ter seus recursos completamente isolados. Em contraste, este artigo apresenta o problema de alocação em nível de aplicações ao invés de VMs ou redes virtuais.

Sendo assim, IoNCloud torna-se uma etapa predecessora utilizada para definir quais aplicações deverão ser alocadas em quais *clusters* virtuais. A abordagem independe do algoritmo posterior utilizado para realizar o posicionamento de VMs dentro de um *data center* [Ballani et al. 2011, Bodík et al. 2012] ou o mapeamento de redes virtuais em substratos de rede [Chowdhury et al. 2009, Alkmim et al. 2011].



(a) VMs dispostas de forma desordenada, compartilhando os recursos de um mesmo ambiente sem isolamento. Tráfegos de diferentes naturezas podem causar comportamento imprevisível.

(b) Agrupamentos de VMs que geram menor entropia (desvio em relação ao comportamento médio) são preferidos. Logo, VMs são atraídas ou repelidas de um mesmo *cluster* virtual de forma a fazer o sistema atingir menor entropia global.

**Figura 2. Desordem na rede da nuvem, uma analogia aos processos naturais.**

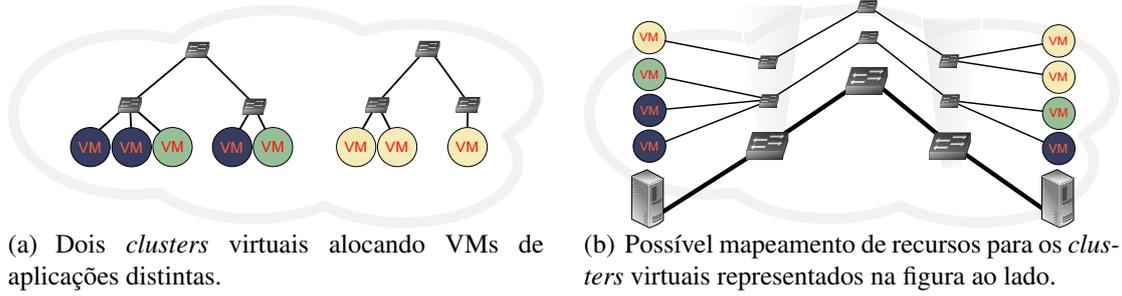
#### 4.1. Clusters Virtuais Orientados a Tráfego

Um *Cluster Virtual* (*Virtual Cluster* – VC) consiste em um conjunto de VMs interligadas através de uma rede virtual, conforme definido em [Xie et al. 2012, Ballani et al. 2011]. Nesse modelo de abstração, cada rede virtual – e, conseqüentemente, cada VC – possui uma parcela de recursos de rede reservados e completamente isolados. Com isso, as VMs de um VC não podem interagir com VMs de qualquer outro VC<sup>2</sup>.

Similarmente aos *clusters* virtuais da literatura, um *Cluster Virtual Orientado a Tráfego* (*Traffic-Oriented Virtual Cluster* – TOVC) é formado por uma topologia de rede virtual genérica (ou seja, não necessariamente limitada a topologias de árvore) conectando um número arbitrário de VMs. A topologia desses *clusters* virtuais é descrita por um grafo, onde os nós virtuais representam máquinas ou *switches* virtuais e as arestas representam enlaces virtuais. Contudo, apesar dos TOVCs de IoNCloud serem fundamentados em propostas anteriores, eles se diferenciam dos tradicionais por uma série de razões, conforme segue:

1. VMs pertencentes a diferentes aplicações e/ou locatários podem compartilhar um mesmo *cluster*. A principal motivação por trás desse compartilhamento é prover uma melhor granularidade na utilização dos recursos de rede.
2. A reserva dos recursos de rede para um determinado TOVC deve garantir a *média agregada dos padrões de tráfego* das VMs das aplicações alocadas. Tal valor pode ser obtido através do histórico da demanda de tráfego das VMs. Essa medida busca prover garantias de tráfego que evitem o desperdício de recursos.
3. O tamanho de cada TOVC (número de aplicações e quantidade de recursos virtuais) varia de acordo com a entropia (ou seja, conforme o grau de interferência entre os tráfegos das aplicações). Em situações limite, o modelo reduz-se às abordagens atuais, onde (i) todas as aplicações são alocadas em um mesmo TOVC ou (ii) cada aplicação é alocada a um TOVC distinto. No primeiro caso não existe isolamento, acarretando imprevisibilidade de desempenho. Em contrapartida, o segundo caso reduz a interferência entre tráfegos ao custo de uma menor utilização de recursos.

<sup>2</sup>VCs protegem apenas o tráfego local, inter-VMs. Existem mecanismos para transpassar o isolamento provido pelos VCs, por exemplo, com o uso de IPs globais (utilizando a Internet). No entanto, tais mecanismos afetam o tráfego externo à rede do *data center* e estão fora do escopo deste trabalho.



**Figura 3. Clusters virtuais do modelo IoNCloud.**

A Figura 3(a) apresenta dois possíveis *clusters* virtuais no modelo IoNCloud. No primeiro *cluster* (esquerda), aplicações distintas compartilham os recursos de rede reservados, enquanto no segundo *cluster* (direita) uma única aplicação utiliza os recursos. A Figura 3(b) demonstra um exemplo simplificado de mapeamento para os recursos virtuais dos *clusters* sobre os dispositivos da rede física. Apesar dos recursos serem alocados sobre os mesmos dispositivos, não existe interferência entre tráfegos de VMs pertencentes a TOVCs distintos.

#### 4.2. Formalização do modelo

Esta subseção formaliza o modelo de agrupamento ótimo de aplicações em TOVCs. Para facilitar a compreensão, a Tabela 1 resume todos os símbolos utilizados na formulação.

O ambiente em consideração é composto por um conjunto de aplicações  $a \in \mathcal{A}$  e suas respectivas máquinas virtuais  $vm \in \mathcal{VM}_a : a \in \mathcal{A}$ . Além disso, considera-se que cada par de VMs  $vm_1, vm_2 \in \mathcal{VM}_a : a \in \mathcal{A}$  possui uma necessidade de comunicação expressa por  $c(vm_1, vm_2) \in \mathbb{B}$ , representando um tráfego  $\pi \in \Pi_a : a \in \mathcal{A}$ . Caso duas VMs precisem se comunicar [ $c(vm_1, vm_2) = 1$ ], a largura de banda média necessária será expressa por  $b(vm_1, vm_2) \in \mathbb{R}^+$  (ou  $b(\pi)$ , onde  $\pi$  é o tráfego agregado entre as mesmas). Nesse contexto, o problema de distribuição de aplicações em TOVCs consiste em um problema de agrupamento ótimo de aplicações  $\mathcal{A}$  sobre um conjunto de TOVCs  $\mathcal{Z}$  que possua entropia  $\mathcal{E}$  mínima, onde  $\mathcal{E}(\Pi)$  denota a entropia de um conjunto de tráfegos.

**Objetivo.** Nesse modelo, a variável  $\alpha_{a,z} \in \mathbb{B}$  indica se a aplicação  $a \in \mathcal{A}$  será alocada ao TOVC  $z \in \mathcal{Z}$ . Dessa forma, a seguinte função expressa o objetivo de otimização da formulação:

$$\min \sum_z \sum_{a_1, a_2}^{\mathcal{A}: a_1 \neq a_2} \alpha_{a_1, z} * \alpha_{a_2, z} * \mathcal{E}(\Pi_{a_1} \cup \Pi_{a_2}). \quad (1)$$

A primeira parte da Equação (1) ( $\sum_z \sum_{a_1, a_2}^{\mathcal{A}: a_1 \neq a_2} \alpha_{a_1, z} * \alpha_{a_2, z}$ ) indica se as aplicações estão alocadas no mesmo TOVC, ou seja, a expressão resulta em 1 se, e somente se, ambas as aplicações são alocadas no mesmo *cluster*. Tal processo é executado para todas as distribuições possíveis de pares de aplicações em TOVCs. A segunda parte da equação [ $\mathcal{E}(\Pi_{a_1} \cup \Pi_{a_2})$ ] calcula a entropia dos tráfegos. Observe que, em virtude da multiplicação entre os dois termos, a entropia só precisa ser calculada para aplicações dentro de um mesmo *cluster*.

**Tabela 1. Lista de símbolos utilizados na formalização do modelo.**

Símbolo	Descrição
$\mathcal{A}$	Conjunto de aplicações.
$\mathcal{VM}_a$	Conjunto de máquinas virtuais de uma aplicação $a \in \mathcal{A}$ .
$\Pi, \Pi_a$	Respectivamente, conjunto de tráfegos quaisquer ( $\Pi$ ) e conjunto de tráfegos pertencentes à uma aplicação $a \in \mathcal{A}$ ( $\Pi_a$ ).
$\mathcal{Z}$	Conjunto de TOVCs.
$\mathcal{E}(\Pi) \in \mathbb{R}^+$	Entropia entre tráfegos pertencentes a um conjunto $\Pi$ .
$c(vm_1, vm_2) \in \mathbb{B}$	Entrada expressando a necessidade de comunicação entre duas VMs quaisquer, onde $vm_1, vm_2 \in \mathcal{VM}_a : a \in \mathcal{A}$ .
$b(vm_1, vm_2) \in \mathbb{R}^+$	Entrada expressando a demanda de banda média necessária para comunicação entre duas VMs, onde $vm_1, vm_2 \in \mathcal{VM}_a : a \in \mathcal{A}$ .
$b(\pi) \in \mathbb{R}^+$	Entrada expressando a demanda de banda média necessária para o tráfego $\pi \in \Pi_a : a \in \mathcal{A}$ .
$\lambda \in \mathbb{R}^+$	Constante definindo um valor máximo para a entropia de um TOVC.
$\alpha_{a,z} \in \mathbb{B}$	Variável binária indicando se a aplicação $a \in \mathcal{A}$ será alocada ao TOVC $z \in \mathcal{Z}$ .
$\delta_{vm,z} \in \mathbb{B}$	Variável binária indicando se a máquina virtual $vm \in \mathcal{VM}_a : a \in \mathcal{A}$ será alocada ao TOVC $z \in \mathcal{Z}$ .
$\gamma^{e^v, \pi} \in \mathbb{B}$	Variável binária indicando se determinado tráfego $\pi$ está utilizando um enlace virtual $e^v$ .
$b(e^v) \in \mathbb{R}^+$	Variável indicando a demanda que deve ser reservada para determinado enlace virtual $e^v$ .

**Restrições.** O processo anterior precisa satisfazer uma série de restrições de forma a tornar o problema válido.

- VMs de uma mesma aplicação precisam ser alocadas no mesmo TOVC, ou seja, se duas VMs quaisquer necessitam se comunicar, então elas devem estar no mesmo *cluster*  $z \in \mathcal{Z}$ . Portanto, seja  $\delta_{vm,z} \in \mathbb{B}$  uma variável binária representando a pertinência da máquina virtual  $vm \in \mathcal{VM}_a : a \in \mathcal{A}$  ao *cluster*  $z \in \mathcal{Z}$ , a Equação (2) precisa ser satisfeita:

$$\delta_{vm,z} = \alpha_{a,z} \quad \forall vm \in \mathcal{VM}_a, \forall a \in \mathcal{A}, z \in \mathcal{Z}. \quad (2)$$

- Aplicações devem ser alocadas a, exatamente, um TOVC. Dessa forma, não pode haver mais do que uma instância de  $\alpha_{a,z}$  com valor 1 *para a mesma aplicação*  $a \in \mathcal{A}$  [Eq. (3)]. No entanto, a recíproca não é verdadeira, uma vez que pode haver diversas aplicações em um mesmo *cluster*  $z \in \mathcal{Z}$  [Eq. (4)].

$$\sum_z \alpha_{a,z} = 1 \quad \forall a \in \mathcal{A}; \quad (3)$$

$$\sum_a \alpha_{a,z} \geq 1 \quad \forall z \in \mathcal{Z}. \quad (4)$$

- A entropia entre dois tráfegos de VMs em um mesmo *cluster*  $z \in \mathcal{Z}$  que não necessitam se comunicar [ $c(vm_1, vm_2) = 0$ ] não deve ser maior do que um limite  $\lambda$  previamente especificado. A constante  $\lambda \in \mathbb{R}^+$ , nesse caso, é um coeficiente não negativo definido pelo provedor de infraestrutura de modo a evitar degradação significativa no desempenho dessas máquinas ao longo do seu período de execução [Eq. (5)].

$$\sum_{a_1, a_2}^A \alpha_{a_1, z} * \alpha_{a_2, z} * [1 - c(vm_1, vm_2)] * \mathcal{E}(\Pi_{a_1} \cup \Pi_{a_2}) \leq \lambda \quad \forall z \in \mathcal{Z}. \quad (5)$$

- Por fim, a função objetivo [Eq. (1)] busca obter a menor entropia possível dentro dos TOVCs, o que levaria no extremo à criação de um TOVC por aplicação. Esse efeito é indesejável, uma vez que acarretaria em baixa utilização de recursos. Portanto, são impostos dois limites, inferior ( $\mathcal{Z}^{inf}$ ) e superior ( $\mathcal{Z}^{sup}$ ), ao número de TOVCs necessários para obter uma solução. Na prática, os valores de  $\mathcal{Z}^{inf}$  e  $\mathcal{Z}^{sup}$  podem ser configurados para os valores limites do modelo (conforme descritos na subseção anterior), ou seja, o valor unitário representa um cenário sem isolamento, enquanto o número máximo de aplicações ( $|\mathcal{A}|$ ) representa isolamento completo. Dessa forma, a primeira tentativa de resolução do problema parte de  $\mathcal{Z}^{inf}$  TOVCs. Caso alguma das restrições anteriores não seja satisfeita, o número de TOVCs é incrementado iterativamente até o limite superior  $\mathcal{Z}^{sup}$ .

Após a distribuição das aplicações em *clusters* virtuais, devem ser utilizados algoritmos de alocação de recursos bem estabelecidos para (i) o posicionamento de VMs [Ballani et al. 2011, Bodík et al. 2012]; e (ii) o mapeamento de redes virtuais em substratos de rede física [Chowdhury et al. 2009, Alkmim et al. 2011]. A única restrição a esse processo deve ser dada na capacidade de transmissão dos enlaces virtuais. Cada enlace virtual  $e^v$  de um cluster  $z \in \mathcal{Z}$  deve possuir uma largura de banda  $b(e^v)$  tal que acomode a soma das demandas médias de todos os seus tráfegos. Dessa forma, seja  $\gamma^{e^v, \pi} \in \mathbb{B}$  a variável que indica se determinado tráfego utiliza o referido enlace virtual. Nesse caso, a seguinte restrição deve ser satisfeita:

$$b(e^v) = \sum_a^{\mathcal{A} : \alpha_{a, z} = 1} \sum_{\pi} \Pi_a b(\pi) * \gamma^{e^v, \pi} \quad \forall e^v \in z, \forall z \in \mathcal{Z}. \quad (6)$$

## 5. Experimentos

A principal contribuição deste artigo é IoNCloud, um modelo amparado nas características de tráfego para oferecer isolamento de recursos às aplicações. Esta seção avalia a abordagem proposta através da análise da (i) interferência entre tráfegos; (ii) fração de utilização dos recursos disponíveis; e (iii) quantidade de banda reservada para cada aplicação e/ou grupo de aplicações. Todos os resultados apresentados utilizam a média de 30 repetições, com intervalo de confiança em 95%.

### 5.1. Cenário de avaliação

Os experimentos realizados seguem um modelo simplificado, baseado no pior caso, onde todas as aplicações de um dado *cluster* virtual compartilham o mesmo enlace virtual.

Seguindo essa premissa, a entropia total do sistema é calculada por meio da análise de um par de VMs de cada aplicação. Ademais, conforme definido na seção anterior, a largura de banda reservada em um enlace virtual é dada pela demanda média do tráfego agregado. Com isso, a interferência é calculada como toda a demanda que ultrapassa essa reserva. Essa situação pode ocorrer mesmo que exista somente uma aplicação por *cluster* virtual.

Para obter a carga de trabalho, desenvolvemos um gerador de tráfegos sintéticos, os quais simulam o comportamento de aplicações típicas de um *data center* (p.ex., HTTP, MapReduce, LDAP etc). Esses tráfegos seguem distribuições com parâmetros conhecidos, conforme definidos por Benson *et. al.* (2010). Ademais, foram utilizadas 12 aplicações simultâneas, igualmente divididas entre 3 tipos de tráfego e agrupadas de acordo com a entropia do tráfego agregado do grupo.

Por fim, o agrupamento de aplicações é restrito às combinações que geram o mesmo número de aplicações por *cluster* virtual. Esse procedimento é necessário pois o cálculo de todas as possíveis combinações torna o processo computacionalmente custoso. Embora tal simplificação acarrete um resultado sub-ótimo, essa primeira análise demonstra que IoNCloud provê isolamento competitivo em relação a abordagens da literatura, a um custo menor na subutilização de recursos.

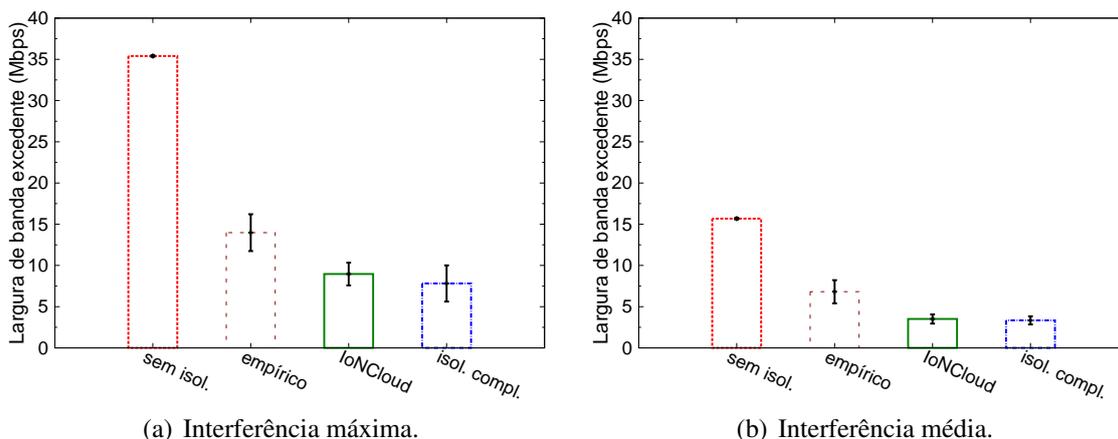
## 5.2. Método de comparação

Em arquiteturas atuais de nuvem pública, sem isolamento de recursos, todos os locatários compartilham os mesmos recursos físicos e estão sujeitos a interferência entre tráfegos. Em resposta, propostas recentes utilizam um *cluster* virtual (ou rede virtual) por locatário [Ballani et al. 2011, Xie et al. 2012] para prover isolamento e garantias de largura banda. Essas abordagens foram denominadas de *sem isolamento*, *isolamento completo*, respectivamente. Em contraste, IoNCloud oferece isolamento compartilhado, permitindo a coexistência das aplicações em um mesmo *cluster* virtual.

A fim de reduzir a interferência, IoNCloud utiliza a entropia (ou seja, desvio em relação à média) como métrica para definir quais aplicações serão agrupadas ou isoladas. Uma variação dessa abordagem consiste em agrupar aplicações em recursos físicos de acordo com métricas variadas, as quais ignoram padrões de tráfego (p.ex, número máximo de VMs compartilhando o mesmo enlace físico [Webb et al. 2011]). Para representar esse tipo de abordagem, aplicações são selecionadas ao acaso para serem agrupadas em *clusters* virtuais. Nos experimentos, esse procedimento foi denominado de *empírico*.

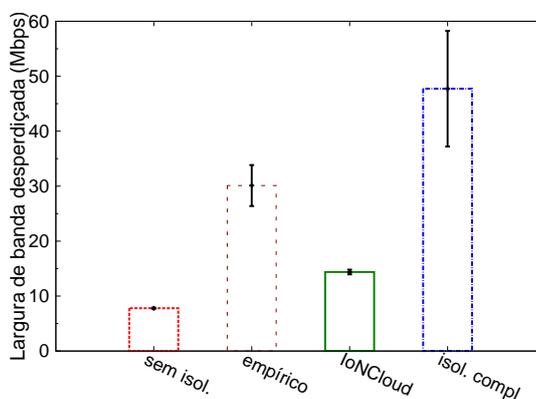
## 5.3. Resultados

A abordagem de isolamento de IoNCloud é competitiva com propostas de isolamento completo. A Figura 4 demonstra a demanda de tráfego máxima e média excedente à reservada. Quando comparado ao caso sem isolamento, IoNCloud consegue reduzir a interferência máxima em aproximadamente 75% (78% na interferência média). Além disso, é possível detectar um ganho no agrupamento baseado na entropia, uma vez que o procedimento empírico comporta-se 36% pior do que IoNCloud para a interferência máxima (49% para a interferência média). Em contraste, quando comparado com o isolamento completo (1 aplicação por TOVC) a diferença para IoNCloud é de apenas 12% na interferência máxima (aproximadamente 5% na interferência média). Esse valor pode tornar-se desprezível quando comparado ao ganho na utilização de recursos, conforme demonstrado na próxima avaliação.



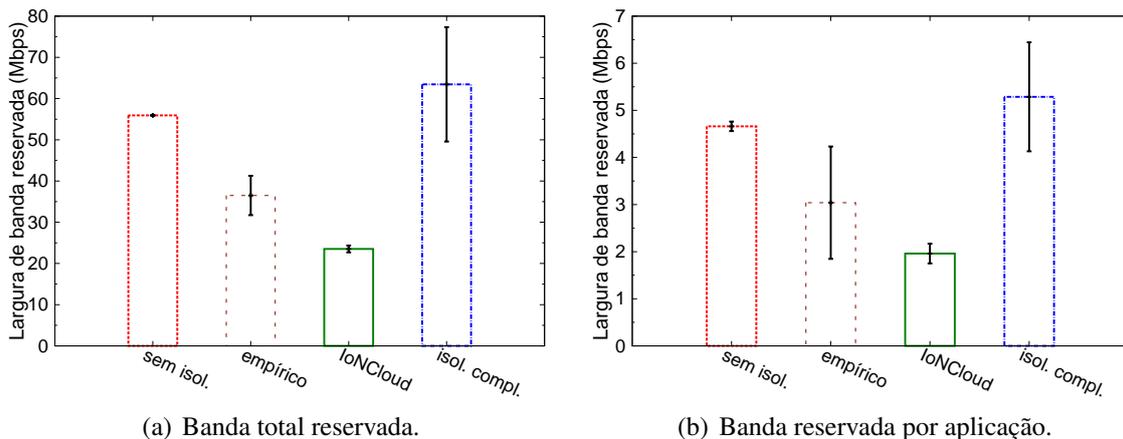
**Figura 4. Interferência entre tráfegos, medida pela banda excedente à reservada.**

*O agrupamento por menor entropia ( $\mathcal{E}$ ) reduz a subutilização absoluta de recursos.* A Figura 5 demonstra a média de recursos ociosos para cada estratégia de isolamento. Como esperado, a subutilização é menor quando não há isolamento. Entretanto, quando comparado com as abordagens que provêm isolamento, a subutilização é menor em IoNCloud, chegando a representar apenas 30% do valor do isolamento completo. Tal fenômeno ocorre porque o agrupamento de tráfegos com menor  $\mathcal{E}$  resulta na menor dispersão em relação à média (vide métrica definida na Seção 3). Logo, a métrica influencia o modelo ao agrupar aplicações cujo tráfego agregado possui poucos picos (valores acima da média – interferência) e vales (valores abaixo da média – subutilização).



**Figura 5. Subutilização média de recursos, medida pela quantidade de recursos ociosos (abaixo do reservado).**

*IoNCloud acarreta menor demanda por reserva de recursos.* A Figura 6 indica a largura de banda total reservada no *data center* da nuvem [Figura 6(a)] e a largura de banda média por aplicação [Figura 6(b)], ou seja, o total reservado para o *cluster* virtual dividido pelo número de aplicações alocadas no *cluster* virtual. Em ambos os casos, IoNCloud reserva uma quantia menor de recursos. Intuitivamente, seria possível inferir que aplicações em IoNCloud são privadas de largura de banda. No entanto, conforme demonstrado nos experimentos, esses recursos acabam tendo uso ineficiente, seja pela interferência dada pela falta de isolamento [Figura 4], ou pelo desperdício [Figura 5], como no caso do isolamento completo. Logo, potenciais benefícios obtidos pela reserva de uma quantidade maior de recursos são apenas aparentes.



**Figura 6. Largura de banda reservada na nuvem.**

## 6. Trabalhos Relacionados

Diversas propostas presentes na literatura buscam oferecer garantias de desempenho às aplicações dos locatários da nuvem. Esta seção discute tanto trabalhos que incluem algum nível de controle de tráfego, quanto trabalhos que efetivamente provêm isolamento de recursos de rede. Alguns mecanismos fundamentais como a limitação da banda de um fluxo (*rate limiting*) [Raghavan et al. 2007] e o posicionamento de VMs (*VM placement*) [Jiang et al. 2012] permeiam a maioria das estratégias de reserva e/ou isolamento presentes nessas propostas, dando sustentação a modelos mais precisos e efetivos.

**Virtualização de redes.** Algumas estratégias recentes fazem uso da virtualização de redes nos *data centers* de nuvem para mitigar a interferência entre tráfegos. Guo *et al.* (2010), Ballani *et al.* (2011) e Rodrigues *et al.* (2011) provêm garantias de banda fixas às comunicações entre pares ou grupos de VMs pertencentes a um mesmo locatário e isoladas em uma rede virtual. Xie *et al.* (2012), por sua vez, exploram a possibilidade das demandas de rede das aplicações variarem ao longo do tempo e propõem um mecanismo de reserva de banda também variante no tempo às VMs isoladas nas redes virtuais. Apesar de oferecerem certas garantias de desempenho, tais abordagens consideram os requisitos de cada locatário individualmente, normalmente resultando em baixa utilização de recursos. Em contraste, Marcon *et al.* (2012) propõem uma estratégia de alocação de recursos baseada no agrupamento de locatários mutuamente confiáveis em domínios virtuais isolados. Tal trabalho é ortogonal ao IoNCloud, uma vez que os autores utilizam confiança mútua como métrica para alocação.

**Compartilhamento proporcional.** Outro grande grupo de propostas envolve o compartilhamento proporcional dos recursos de rede entre as VMs dos locatários. Lam *et al.* (2012) adotam a atribuição de pesos específicos para cada locatário da nuvem para, então, alocar os recursos de enlaces congestionados com base nesses pesos. Por sua vez, Popa *et al.* (2012) propõem mecanismos eficientes para a utilização proporcional dos recursos de rede na presença de garantias mínimas de banda. Essas propostas, apesar de facilitarem a previsibilidade de desempenho das aplicações, possuem alto custo de gerenciamento de recursos de rede.

**Limitação no uso de recursos.** Algumas alternativas envolvem a limitação do número de VMs e locatários que compartilham os recursos de um determinado enlace ou porção da rede. A Amazon (2012), por exemplo, oferece serviços que reservam uma porção dos recursos da rede para a comunicação exclusiva entre as VMs de um locatário. De forma mais relaxada, a estratégia de Webb *et. al.* (2011) permite que os locatários especifiquem o número de tarefas (VMs) que compartilham um determinado enlace. Esse grupo de estratégias normalmente aumenta os custos dos locatários e diminui a eficiência na utilização dos recursos dos *data centers*, uma vez que o tráfego da maioria das aplicações de nuvem é de natureza intermitente.

**Escalonamento de tarefas.** Lago *et. al.* (2012) propõem o escalonamento de tarefas ciente do consumo de energia nas máquinas virtuais da nuvem. Apesar dos benefícios relacionados ao controle de energia, a proposta não trata da interferência de desempenho causada pela heterogeneidade do tráfego presente na rede interna da nuvem.

## 7. Considerações Finais

Embora a computação em nuvem traga diversas vantagens, o compartilhamento dos recursos de rede entre um elevado número de aplicações, potencialmente com requisitos de largura de banda distintos, acarreta uma variabilidade significativa no desempenho de rede das aplicações. Consequentemente, o tempo de execução dessas aplicações é negativamente afetado e o custo dos locatários torna-se imprevisível.

Neste artigo, foi proposto um modelo de reserva e isolamento de recursos de rede da nuvem (IoNCloud). Nesse modelo, instâncias de aplicações são agrupadas em *clusters* virtuais orientados a tráfego (TOVCs), com o objetivo de isolar aplicações cujo compartilhamento dos recursos de rede ocasiona imprevisibilidade de desempenho. Resultados demonstram que IoNCloud provê dois benefícios chave. Por um lado, locatários obtêm melhor previsibilidade de desempenho, uma vez que IoNCloud é competitivo em relação a estratégias de isolamento completo. Por outro, provedores de infraestrutura obtêm melhor multiplexação de recursos, visto que IoNCloud apresenta subutilização de recursos 70% menor do que o isolamento completo.

Nos trabalhos em andamento, estamos desenvolvendo heurísticas para resolver o problema de forma eficiente, viabilizando a aplicação da proposta em nuvens de larga escala. Em trabalhos futuros, vislumbramos novos algoritmos de posicionamento de VMs que possam explorar melhor as características dos TOVCs. Também estuda-se a possibilidade de agregar maior dinamicidade ao ambiente, permitindo que aplicações e *clusters* sejam migrados para atender novas demandas de tráfego.

## Referências

- [Abts and Felderman 2012] Abts, D. and Felderman, B. (2012). A guided tour of data-center networking. *Commun. ACM*, 55(6):44–51.
- [Alkmim et al. 2011] Alkmim, G., Batista, D. M., and da Fonseca, N. L. S. (2011). Mapeamento de redes virtuais em substratos de rede. In *Anais do XXIX SBRC*.
- [Amazon 2012] Amazon (2012). High performance computing (hpc) on aws. Disponível em <https://aws.amazon.com/hpc-applications/>.

- [Armbrust et al. 2010] Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., Stoica, I., and Zaharia, M. (2010). A view of cloud computing. *Commun. ACM*, 53(4):50–58.
- [Ballani et al. 2011] Ballani, H., Costa, P., Karagiannis, T., and Rowstron, A. (2011). Towards predictable datacenter networks. In *Proceedings of the ACM SIGCOMM '11 conference*.
- [Benson et al. 2010] Benson, T., Akella, A., and Maltz, D. A. (2010). Network traffic characteristics of data centers in the wild. In *Proceedings of the 10th ACM IMC conference*.
- [Bodík et al. 2012] Bodík, P., Menache, I., Chowdhury, M., Mani, P., Maltz, D. A., and Stoica, I. (2012). Surviving failures in bandwidth-constrained datacenters. In *Proceedings of the ACM SIGCOMM '12 conference*.
- [Chowdhury et al. 2009] Chowdhury, N., Rahman, M., and Boutaba, R. (2009). Virtual network embedding with coordinated node and link mapping. In *Proceedings of the 28th IEEE INFOCOM conference*.
- [Genez et al. 2012] Genez, T. A. L., Bittencourt, L. F., and Madeira, E. R. M. (2012). Discretização do tempo na utilização de programação linear para o problema de escalonamento de workflows em múltiplos provedores de nuvem. In *Anais do XXX SBRC*.
- [Guo et al. 2010] Guo, C., Lu, G., Wang, H. J., Yang, S., Kong, C., Sun, P., Wu, W., and Zhang, Y. (2010). Secondnet: a data center network virtualization architecture with bandwidth guarantees. In *Proceedings of the 6th ACM Co-NEXT conference*.
- [Jiang et al. 2012] Jiang, J., Lan, T., Ha, S., Chen, M., and Chiang, M. (2012). Joint vm placement and routing for data center traffic engineering. In *Proceedings of the 31st IEEE INFOCOM conference*.
- [Kandula et al. 2009] Kandula, S., Sengupta, S., Greenberg, A., Patel, P., and Chaiken, R. (2009). The nature of data center traffic: measurements & analysis. In *Proceedings of the 9th ACM IMC conference*.
- [Lago et al. 2012] Lago, D. G., Madeira, E. R. M., and Bittencourt, L. F. (2012). Escalonamento com prioridade na alocação ciente de energia de máquinas virtuais em nuvens. In *Anais do XXX SBRC*.
- [Lam et al. 2012] Lam, V. T., Radhakrishnan, S., Pan, R., Vahdat, A., and Varghese, G. (2012). Netshare and stochastic netshare: predictable bandwidth allocation for data centers. *ACM SIGCOMM Comput. Commun. Rev.*, 42(3):5–11.
- [Marcon et al. 2012] Marcon, D. S., Neves, M. C., Oliveira, R. R., Buriol, L. S., Gaspary, L. P., and Barcellos, M. P. (2012). Mitigando ataques de egoísmo e negação de serviço em nuvens via agrupamento de aplicações. In *Anais do XII SBSeg*.
- [Popa et al. 2012] Popa, L., Kumar, G., Chowdhury, M., Krishnamurthy, A., Ratnasamy, S., and Stoica, I. (2012). Faircloud: sharing the network in cloud computing. In *Proceedings of the ACM SIGCOMM '12 conference*.
- [Raghavan et al. 2007] Raghavan, B., Vishwanath, K., Ramabhadran, S., Yocum, K., and Snoeren, A. C. (2007). Cloud control with distributed rate limiting. In *Proceedings of the ACM SIGCOMM '07 conference*.
- [Rodrigues et al. 2011] Rodrigues, H., Santos, J. R., Turner, Y., Soares, P., and Guedes, D. (2011). Gatekeeper: supporting bandwidth guarantees for multi-tenant datacenter networks. In *Proceedings of the 3rd USENIX WIOV workshop*.
- [Webb et al. 2011] Webb, K. C., Snoeren, A. C., and Yocum, K. (2011). Topology switching for data center networks. In *Proceedings of the 11th USENIX Hot-ICE conference*.
- [Xie et al. 2012] Xie, D., Ding, N., Hu, Y. C., and Kompella, R. (2012). The only constant is change: incorporating time-varying network reservations in data centers. In *Proceedings of the ACM SIGCOMM '12 conference*.