

# Um Sistema Multinível de Distribuição de Identidades em Névoas Computacionais

Bruno Cremonezi<sup>1</sup>, Alex Vieira<sup>2</sup>, José Nacif<sup>3</sup>, Michele Nogueira<sup>1</sup>, Aldri Santos<sup>1</sup>

<sup>1</sup>Depto. de Informática – Universidade Federal do Paraná (UFPR)

<sup>2</sup>Depto. de Ciência da Computação – Universidade Federal de Juiz de Fora (UFJF)

<sup>3</sup>Instituto de Ciências Exatas e Tecnológicas – Universidade Federal de Viçosa (UFV)

{bmcremonezi,michele,aldri}@inf.ufpr.br, alex@ufjf.edu.br, jnacif@ufv.br

**Abstract.** Fog computing operates together with traditional clouds to deliver more efficiently services. The fogs replicate services and data in structures closer to users with the goal of increasing the efficiency of services offered. However, such replication can bring problems related to the privacy and data vulnerability. For example, identity management and storage is a service that must be well evaluated before being indiscriminately replicated in fogs. In fact, this type of service was initially proposed as a cloud-centric service. Thus, identity replication is not a trivial problem, because current identity replication techniques generate either indiscriminate replication or little latency reduction. In this study, we investigated the relationship between replica number and the user's distance in identity distribution systems with probabilistic replication. The simulation results show that, when compared to a well-known distribution systems presented in the literature, the probabilistic distribution reduced the number of jumps need to recover the identities in 25% and in 30% the number of replicas along the network.

**Resumo.** As névoas computacionais atuam em conjunto com as tradicionais nuvens para entregar serviços de forma mais eficiente. As névoas replicam os serviços e os dados em estruturas mais próximas aos usuários com o objetivo de aumentar a eficiência de serviços ofertados. No entanto, tal replicação traz problemas relacionados a, por exemplo, serviços de gestão e o armazenamento de identidades. De fato, este tipo de serviço foi proposto como um serviço centralizado nas nuvens. Assim, a replicação de identidades não é um problema trivial, pois as atuais técnicas de replicação de identidades ou geram uma replicação indiscriminada ou pouco reduzem a latência de recuperação dessas. Neste estudo, nós investigamos a relação entre número de réplicas e a distância de seu usuário em sistemas de distribuição de identidades com replicação probabilística. Os resultados de simulação mostram que, quando comparado a outros sistemas de distribuição, a distribuição probabilística reduz em 25% o número de saltos necessários para recuperar a identidade e em 30% o número de réplicas pela rede.

## 1. Introdução

Estima-se que o número de dispositivos portáteis conectados a Internet excede o número de pessoas no planeta. De acordo com a Cisco Systems, atualmente existem, em média, 1,4 dispositivos móveis por pessoa e este número tende a crescer ainda mais [Cisco 2016].

Desta forma, a demanda por uma identidade virtual (ID) se torna um problema cada vez mais complexo. Com a Internet oferecendo aplicações que pretendem ser universais e onipresentes [Paul et al. 2011], as IDs desempenham um papel fundamental nesse contexto. Através delas é possível aplicar medidas de gerenciamento de acesso, autenticação e prestação de serviço [Torres et al. 2013, Neto et al. 2016]. Neste contexto, o gerenciamento de identidades (*Identity Management* ou IdM) atrai atenção como uma forma eficiente de prover confiança entre as entidades [Sabena et al. 2010].

A identificação como serviço (*Identity as a Service* ou IDaaS) é um IdM na nuvem, fora dos limites da organização e dos aplicativos que os utilizam. O serviço de IdM passa a ser terceirizado e a nuvem gerencia as identidades de usuários externos. Essa arquitetura permite um gerenciamento centralizado de diversas identidades de dispositivos, no qual toda requisição de identidade é direcionada para as IDaaS na nuvem. Além de fornecer as funções básicas de gerenciamento de identidades, a IDaaS oferece os benefícios que a arquitetura de nuvem traz consigo como um baixo custo de hardware, gerenciamento fácil e alta gama de opções de integração [Habiba et al. 2014]. No entanto, em aplicações sensíveis ao atraso, a IDaaS pode não ser a solução mais adequada. A natureza centralizada da nuvem e a distância entre esta e o dispositivo que requisita a identidade impede seu uso nestes casos [Gomez-Cárdenas et al. 2018].

A computação em névoa é um paradigma proposto inicialmente pela Cisco [Bonomi et al. 2012] que estende o conceito de nuvem para os dispositivos de borda da rede. Com capacidade de ser utilizada de maneira autossuficiente ou cooperativa com as nuvens, a névoa é capaz de oferecer serviços próximos aos usuários, o que possibilita a execução de serviços com uma latência menor e um menor tráfego de rede [Abbas et al. 2018]. Na IDaaS, isso significa que a névoa será capaz utilizar os dispositivos intermediários da rede para o armazenamento e a gerenciamento das identidades. No entanto, devido à natureza distribuída da computação de névoa, acomodar uma IDaaS nesta arquitetura requer um gerenciamento de identidades mais dinâmico e robusto [Dsouza et al. 2014].

Em um contexto cooperativo entre nuvem-névoa, as identidades permanecem armazenadas na nuvem, porém com diversas réplicas distribuídas em pontos estratégicos na névoa. Logo, pode-se dizer que o conceito de centralização das identidades será substituído por um modelo distribuído com o objetivo de fornecer uma resposta mais rápida da requisição de identidade se comparada com uma IDaaS na nuvem. Em uma IDaaS na névoa, toda requisição de identidade é inicialmente solicitada para a névoa. Caso esta já possua uma réplica da identidade, a requisição é respondida de maneira direta. Caso contrário, a névoa encaminha essa solicitação para a nuvem, que garante a resposta da requisição da identidade com latência correspondente ao acesso na nuvem. Nesse cenário, determinar onde e como replicar as identidades representa um desafio de pesquisa relevante, uma vez que o posicionamento adequado das réplicas das identidades é de grande importância para o sucesso uma IDaaS na névoa.

Um sistema de distribuição de identidades em névoas consiste de: um algoritmo de replicação de identidades e uma política de substituição de identidades. O primeiro seleciona onde as identidades serão armazenadas e o segundo descarta identidades menos utilizadas com o intuito de armazenar identidades de maior interesse pela rede. Em essência, o número de réplicas geradas por esse sistema acaba por comprometer a eficiência do IDaaS na Névoa. Uma distribuição com poucas réplicas resulta em um alto número

de requisições à nuvem, o que remove todos os benefícios da névoa [Gomez-Cárdenas et al. 2018]. Uma replicação indiscriminada das identidades na névoa é custosa e subótima [Chai et al. 2012]. Além disso, estratégias sofisticadas de replicação, em geral, adicionam complexidade e custos computacionais que podem não ser suportados pelos dispositivos da névoa, tornando-as não aplicáveis do ponto de vista prático. Os sistemas de replicação probabilísticos se apresentam como uma solução simples, e com baixo custo computacional, para a replicação das identidades em névoas. Nesses sistemas, os dispositivos da névoa possuem uma determinada probabilidade de gerar e armazenar uma réplica das identidades. Essa solução é capaz de superar problemas de replicação indiscriminada e complexidade, pois oferece uma baixa redundância e um número aceitável de replicações de identidades na rede.

Tendo em mente esse contexto, o presente trabalho tem como objetivo avaliar a qualidade de diferentes sistemas de distribuição se comparado ao sistema de distribuição de identidades probabilísticos. Com este trabalho, verificaremos se com uma simples replicação probabilística, o sistema é capaz de ser mais eficaz que tradicionais sistemas de replicação. A avaliação da distribuição probabilística ocorreu através simulação de um IDaaS em névoa em que diferentes políticas de substituição de identidades foram comparadas com o algoritmo de replicação probabilístico, sendo elas: “menos recentemente utilizado” (*Least Recently Used*, LRU); “primeiro a entrar e primeiro a sair” (*First In, First Out*, FIFO) e a “substituição aleatória” (*Randomly Replace*, RR). O cenário de avaliação consiste em uma topologia virtual de árvore, em que diversos dispositivos requisitam as identidades [Roman et al. 2018]. Com o objetivo de avaliar o comportamento do sistema em diferentes cenários, foram atribuídas diferentes capacidades de armazenamento nos dispositivos da névoa e diferentes popularidades das identidades, desde cenários com abundância até cenários com escassez de recursos. Assim, o desempenho da distribuição probabilística foi avaliado em relação às demais distribuições. Foi verificado o balanço entre o número de réplicas e o tempo de recuperação das identidades. Os resultados mostram que a distribuição de identidades com replicação probabilística oferece uma economia de até 25% no número médio de saltos necessários para a recuperação das identidades. Esta distribuição foi capaz de oferecer uma redução de até 30% do número de réplicas de identidades na névoa ao ser comparado com abordagens clássicas.

A organização do artigo é descrita a seguir. A Seção 2 discute os trabalhos relacionados. A Seção 3 detalha o modelo de sistema. A Seção 4 detalha os sistemas de distribuição das identidades em névoa apresentados nesse trabalho. A Seção 5 apresenta a metodologia de avaliação de desempenho. A Seção 6 apresenta os resultados. Finalmente, a Seção 7 conclui o artigo e apresenta direções futuras.

## **2. Trabalhos Relacionados**

O trabalho de [Trnka and Cerny 2016] propõe uma estrutura de gerenciamento de identidades para o ambiente de nuvens computacionais que armazena as identidades de dispositivos em um banco de dados centralizado. Nessa proposta, a fim de iniciar a utilização de algum serviço, o dispositivo solicitante deve recuperar sua identidade através de um gerenciador de identidades. Esse gerenciador, ao receber essa solicitação, recupera a identidade do dispositivo no banco de dados centralizado e a encaminhando-as para o solicitante. O principal problema dessa proposta é o ponto único de falha devido à natureza centralizada tanto do banco de dados quanto do gerenciador de identidades. Caso ocorra

uma falha no banco de dados ou no gerenciador, conseqüentemente, todo sistema que implementa esta solução para o gerenciamento de identidades irá falhar também.

Um dos trabalhos pioneiros que propõe um sistema de gerenciamento de identidade em uma arquitetura integrada de nuvem e névoa é proposto em [Gomez-Cárdenas et al. 2018]. Esse trabalho propõe um sistema para gerenciar identidades desde a nuvem até as bordas da rede. Esse sistema fornece benefícios importantes para os usuários finais, como a possibilidade de executar serviços com uma latência reduzida e suporte aos nós em movimento. Além disso, essa proposta fragmenta as identidades em pequenos pedaços, o que aprimora o controle que os usuários têm sobre seus dados, e utiliza dispositivos agregadores para a recuperação das identidades na névoa. Os resultados apresentados mostraram que a névoa foi capaz de reduzir o atraso do acesso a identidade, porém o desempenho está altamente relacionado com o posicionamento adequado dos fragmentos das identidades na névoa.

No contexto de distribuição de identidades em névoa, a replicação das identidades é uma área ainda não explorada. No entanto, a avaliação de esquemas de armazenamento distribuído e políticas de substituição de conteúdo tem sido extensivamente estudada no contexto de redes centrada em conteúdo (CNN). Diversos esquemas de cache baseados em popularidade de conteúdo foram apresentados [Wu et al. 2013, Bernardini et al. 2013]. Porém, apesar de eficientes, esses esquemas necessitam de uma sincronização entre os nós, introduzindo uma grande complexidade tanto de rede quanto computacional. Sendo assim, tais abordagens não são aplicáveis em nosso contexto.

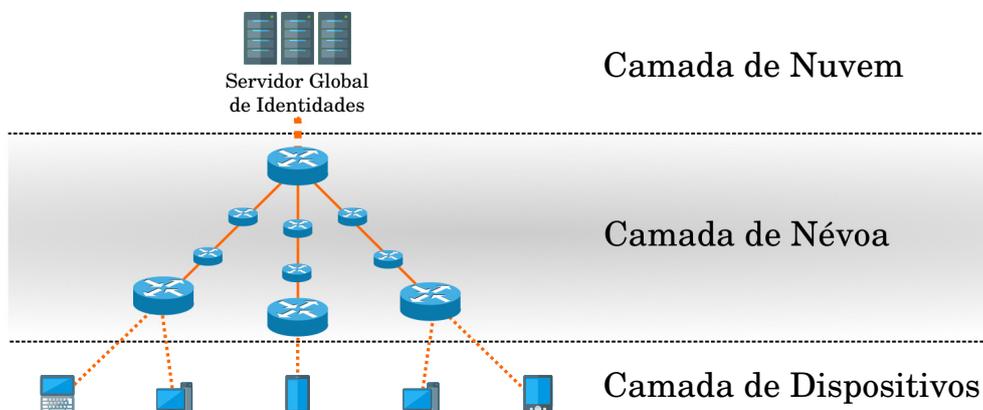
Em [Rossi and Rossini 2011] foram avaliadas diferentes combinações de esquemas de cache e políticas de substituição de conteúdo. Tal estudo simulou diversas topologias de rede e os resultados mostraram que a eficiência de uma política de replicação aleatória é superior a outras políticas de replicação. No entanto, pouco foi discutido sobre o comportamento de seus resultados.

O presente trabalho avança no estado da arte na medida em que propõe o uso de políticas de substituição em cache para replicação de identidades em névoas computacionais. Discute-se sobre o comportamento do modelo probabilístico nesse novo contexto e advoga-se que o modelo probabilístico é aplicável no contexto de replicação de identidades em névoas, pois permite uma distribuição rápida das identidades com uma considerável redução do número de réplicas e acesso à nuvem.

### **3. Modelo do Sistema**

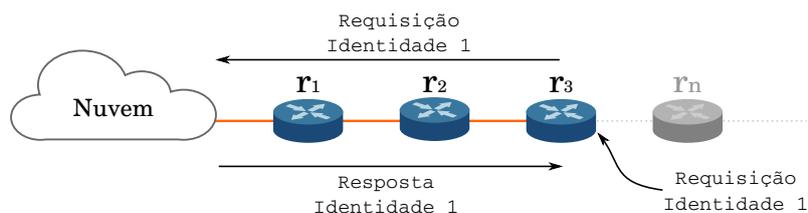
A topologia da plataforma de computação em névoa considerada neste trabalho pode ser vista na Figura 1. Nela, três camadas hierárquicas são definidas: a camada de Nuvem, camada de Névoa e camada de Dispositivos. A camada de Nuvem contém o Servidor Global de Identidades, capaz de fornecer a identidade de qualquer um dos  $N$  dispositivos da camada de Dispositivos. A camada de névoa, foco deste trabalho, é composta por múltiplas redes [Roman et al. 2018]. Essas redes são capazes de estabelecer uma comunicação com a nuvem e armazenar, de maneira temporária, até  $n$  (com  $n < N$ ) identidades requisitadas pelos dispositivos da camada de dispositivos. É importante notar que a névoa pode ser constituída por múltiplos níveis hierárquicos e é composta de diversos tipos de dispositivos com menor capacidade computacional se comparado com a Nuvem, como pontos de acesso, estações base, terminais e etc. Todo dispositivo da camada de dispo-

sitivo não possui acesso direto a Nuvem, sendo assim eles são conectados em ao menos uma rede da névoa e realiza requisições de identidade.



**Figura 1. Topologia da plataforma de computação em Névoa**

As requisições de identidade são direcionadas para a nuvem em um caminho composto por  $r$  redes na névoa. A Figura 2 exemplifica uma requisição em um caminho composto por 3 redes na névoa. Caso alguma dessas redes da névoa possua a identidade requisitada pelo dispositivo, o encaminhamento da requisição de identidade para o servidor de identidade é interrompido e a resposta da solicitação parte desta rede da névoa. Como a topologia da rede é hierárquica, todas mensagens de requisição e resposta das identidades seguem a mesma rota [Psaras et al. 2012]. Durante a resposta de identidade, uma rede da névoa ao receber a identidade a encaminha para o usuário e segue uma política de replicação que determina, com apoio de um conjunto de regras preestabelecidas, se a rede da névoa vai gerar e armazenar uma réplica dessa identidade.



**Figura 2. Requisição e resposta de identidade à nuvem**

#### 4. Distribuição das identidades em Névoa

Para a distribuição das identidades na Névoa, são necessários: o algoritmo de replicação de identidades e uma política de substituição de identidades. O primeiro algoritmo determina se alguma das  $r$  redes da névoa presentes no caminho da resposta de identidade irá gerar e armazenar uma réplica da identidade transitada. Caso a capacidade de armazenamento da rede da névoa não esteja cheia, a identidade é replicada e armazenada para futuros usos. No entanto, em casos nos quais a capacidade de armazenamento das identidades está cheia, uma política de substituição das identidades é necessária para determinar qual identidade será descartada com o objetivo de liberar espaço para a nova réplica de

identidade. De maneira geral, o objetivo da política de substituição de identidades é manter as identidades que mais provavelmente serão requisitadas e descartar as que tendem a ser menos requisitadas no futuro.

#### 4.1. Políticas de replicação de identidades

Neste trabalho, assume-se que todas as redes da névoa devem executar individualmente e de maneira distribuída as políticas de replicação de identidades. Essa suposição pretende explorar os benefícios da nuvem, em que as redes da névoa, ainda que conectadas, não possuem alto poder computacional. Portanto, concentramos nas políticas de replicação mais simples e comumente conhecidas: replicação universal (LCE), replicação de um nível (LCD) e a replicação probabilística (Prob).

A primeira política de replicação universal baseia-se em *Leave a Everywhere* (LCE) [Jacobson et al. 2009]. No contexto de Névoa, esta política de replicação de identidades sugere que todo roteador do caminho deve armazenar a identidade de resposta. Por sempre gerar e armazenar uma réplica da identidade, essa política distribui rapidamente as identidades na Névoa. No entanto, colocar réplicas de uma mesma identidade em vários pontos da névoa é altamente custoso para seus dispositivos e gera muitas redundâncias. Na Figura 2, caso a replicação universal seja aplicada, todas as redes da névoa do caminho da requisição, i.e., os roteadores  $r_1$ ,  $r_2$  e  $r_3$ , irão replicar a Identidade.

A política de replicação de um nível baseia-se em *Leave a Copy Down* (LCD) [Lautaris et al. 2006]. Esta política de replicação de identidades sugere que ao encontrar uma identidade, somente o roteador em um nível abaixo deve replicar e armazenar a identidade transitada. Considerada uma abordagem mais conservadora no quesito replicação, tal política permite que identidades mais requisitadas se aproximem da borda da rede, enquanto as menos requisitadas permanecem nos nós mais distantes da Névoa. Na Figura 2, caso a replicação de um nível seja aplicada, somente a rede da névoa logo abaixo da nuvem (roteador  $r_1$ ) replica a identidade 1. Note que, caso  $r_3$  receba novamente uma requisição de identidades, a identidade será encontrada em  $r_1$  e uma réplica da identidade será criada em  $r_2$ , e assim sucessivamente.

A política de replicação probabilística [Psaras et al. 2012], conhecida como *Prob(p)*, armazena a identidade em cada roteador na rota de resposta com uma probabilidade  $p$ , com  $0 < p < 1$ . É importante notar que a Replicação Universal é um caso especial da replicação probabilística, com  $p = 1$ . Ao assumir um  $p$  próximo de 1, a replicação Probabilística tende a ter comportamentos similares a Replicação universal: alta distribuição de identidades, porém com alta redundância. Para que ocorra uma diversidade maior, decrementar o valor de  $p$  é necessário. No entanto, definir um valor pequeno de  $p$  pode resultar em uma demorada distribuição das identidades. Essa distribuição lenta pode levar ao desempenho ruim dessa política, pois resulta em diversas requisições para a nuvem. Na Figura 2, caso a replicação probabilística seja aplicada, todas as redes da névoa do caminho da requisição irão replicar a identidade com uma probabilidade  $p$ .

#### 4.2. Políticas de substituição de identidades

Como a capacidade dos dispositivos da névoa é reduzida, cada rede da névoa não é capaz de armazenar todas as identidades da nuvem simultaneamente. Sendo assim, se as capacidades de armazenamento estiverem cheias, elas devem descartar alguma identidade

antes de gerar e armazenar uma réplica da nova identidade. A política de substituição determina qual identidade irá ser descartada. Nesse trabalho exploramos substituições que possuam complexidade  $\mathcal{O}(1)$  para o descarte, foram: Menos recentemente utilizada (*Least Recently Used*, LRU), em que a identidade que foi utilizada menos recentemente é descartada; Primeiro a entrar e primeiro a sair (*First In, First Out*, FIFO), em que a primeira identidade armazenada é descartada; Substituição aleatória (*Randomly Replace*, RR), onde uma identidade qualquer é descartada.

## 5. Avaliação

Como parte da avaliação do comportamento de sistemas de replicação das identidades em Névoa, foram simulados e comparados vários sistemas clássicos de substituição. Para isso, foi construído um simulador de redes orientado a eventos que geram requisições de identidade de acordo com uma distribuição Zipf [Adamic 2000]. Dado que o objetivo final deste trabalho é avaliar a eficiência de tais sistemas através do melhor posicionamento das identidades e a redução de redundância das identidades nas redes da névoa, as métricas mais diretas para avaliar este sistema são: a distância média de encontro da identidade, número de requisições para nuvem e o número de réplicas no sistema. A distância média de encontro das identidades é mensurada pelo número de saltos necessários para encontrar uma rede da névoa que possua a identidade ou o servidor global das identidades. Através dela é possível mensurar os ganhos de se utilizar identidades distribuídas pelo sistema. O número de requisições na nuvem determina que, caso uma requisição de identidade seja encontrada somente no servidor global de identidades, o número de médio de saltos se eleva, conseqüentemente se reduz os ganhos do sistema. O número de réplicas em redes da névoa, por sua vez, representa a redundância das identidades. A criação de um elevado número réplicas indica que as redes da névoa apresentam alta redundância.

A topologia do sistema foi inspirada no trabalho [Psaras et al. 2012] e segue uma árvore com 143 nós no total. A disposição do servidor global de identidades e das redes da névoa segue uma topologia de árvore binária de 5 níveis, totalizando 63 nós. O nó raiz representa o servidor global das identidades com capacidade de armazenamento de até 500 identidades enquanto os demais nós representam a Névoa. Nesta camada, temos diversas redes da névoa conectadas com capacidade reduzida de armazenamento das identidades. Assume-se que a rede não possui dispositivos conectados nas redes da névoa centrais, sendo assim, somente as redes da névoa dos últimos dois níveis recebem requisições de identidades dos dispositivos. Especificamente, cada rede da névoa no último nível desta árvore possui 2 dispositivos conectados, enquanto no penúltimo nível cada rede da névoa possui um dispositivo conectado. Ao total, 80 dispositivos requisitam identidades.

Nas simulações, a capacidade de armazenamento de cada rede da névoa variou entre 1%, 2%, 5%, 10%, 25% e 50% da capacidade total do servidor global de identidades, com o objetivo de avaliar a capacidade do sistema em cenários com abundância e escassez de recursos. Caso seja necessário substituir uma identidade na rede da névoa, os algoritmos *Least Recently Used* (LRU), *First In, First Out* (FIFO) e *Randomly Replace* (RR) foram utilizados para excluir as identidades. Cada um dos dispositivos gera 200 requisições, o que totaliza em 16000 requisições. Essas requisições seguem a distribuição de Zipf com expoente  $\alpha = 0,9$ , com o objetivo de capturar cenários em que as frequências de requisição das identidades são variadas. Em toda simulação, o desempenho dos sistemas de replicação foi avaliado em relação a política: *i) Leave a Copy Everywhere* (LCE),

ii) *Leave Copy Down* (LCD) e o iii) Probabilístico  $p$ . Nesse trabalho, dois valores de  $p$  foram avaliados:  $p = 0,3$  e  $p = 0,7$ .

## 6. Resultados

Esta seção apresenta a avaliação de sistemas de distribuição das identidades em névoa. Essencialmente, busca-se responder as perguntas de pesquisa: (i) Qual número de réplicas geradas por esses sistemas? (ii) Qual o percentual de requisições de identidades encaminhadas à nuvem? (iii) Qual a relação entre número de réplicas de requisições para nuvem? (iv) Quais os ganhos, em termos de redução de saltos, que os sistemas fornecem?

### 6.1. Número de Réplicas

A Figura 3 ilustra um comparativo do número de réplicas geradas em relação ao algoritmo de replicação e política de substituição utilizada. É possível observar que o comportamento do número de réplicas geradas é influenciado principalmente pelo algoritmo de replicação. Dentre os algoritmos de replicação de identidade, foi possível observar que os algoritmos LCE e Prob( $p=0,7$ ), obtiveram uma diferença fixa de 30% do número de identidades. Além disso, ao serem utilizados em redes da névoa com capacidade de armazenamento igual 1% do servidor global de identidades, replicaram aproximadamente 5 vezes mais de identidades em relação ao LCD e Prob( $p=0,3$ ), independente da política de substituição. No entanto, à medida que se aumenta a capacidade da rede da névoa, o número de identidades geradas por tais algoritmos tende a diminuir. Ao serem utilizadas em redes da névoa com capacidade de armazenamento igual 50% do servidor global de identidades, a diferença do número de réplicas geradas pelo LCE e Prob( $p=0,7$ ) em relação ao LCD e Prob( $p=0,3$ ) decai até 1,5 vezes.

Apesar de não ser o maior fator de variação para o número de identidades geradas, a Figura 3 mostra que a medida em que se aumenta a capacidade de armazenamento da rede da névoa o LRU se sobressai em relação ao FIFO e RR. É possível perceber que todos algoritmos de replicação ao serem utilizados em conjunto com o LRU em redes da névoa com capacidade de 25% e 50% de armazenamento, geraram em média 50% menos identidades que se utilizados com o FIFO e RR. No entanto, em redes da névoa com capacidade menor de armazenamento, a diferença foi mínima.

Complementando a análise recém apresentada, destaca-se, ainda, que o alto número de réplicas apresentado pelo LCE e Prob( $p=0,7$ ), justificada principalmente pela alta taxa de replicação desses algoritmos, exige com uma frequência maior o descarte de identidades nas redes da névoa. Nesse cenário, quando a capacidade de armazenamento é baixa, independente do algoritmo de substituição, uma identidade popular tem uma maior chance de ser descartada, justificando o mesmo comportamento para todas as políticas de substituição. No entanto, à medida que se aumenta a capacidade de armazenamento desses dispositivos, observa-se uma diminuição significativa do número de substituições. Nota-se que a política LRU prioriza manter identidades mais populares, consequentemente, observa-se um menor número de réplicas distribuídas na rede.

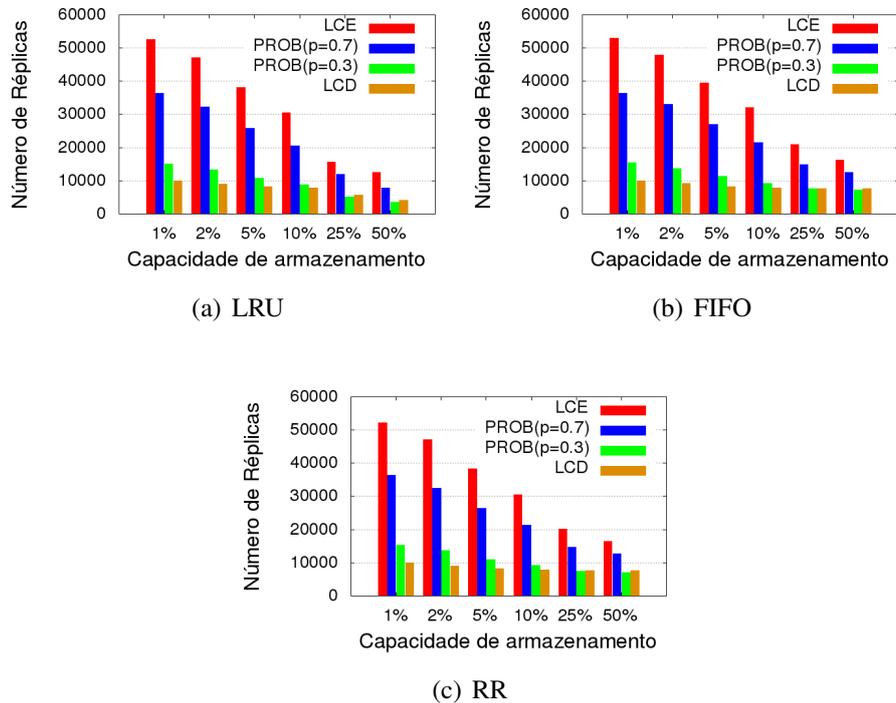
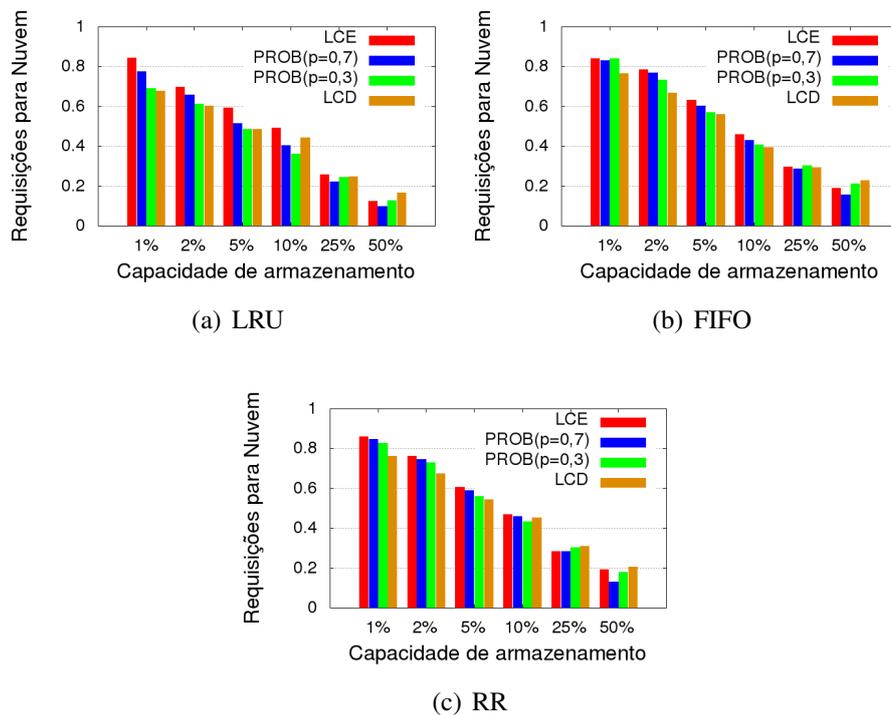


Figura 3. Número de réplicas em relação a política de substituição de identidades

## 6.2. Número de requisições para nuvem

A Figura 4 mostra o número de requisições necessárias em relação à capacidade da armazenagem da rede da névoa política de substituição e algoritmo de replicação. É possível observar que em média, a política de substituição LRU apresentou melhores resultados em relação aos demais. Esse resultado é evidenciado a medida que se aumenta a capacidade de armazenagem da névoa. No entanto, independente da política de substituição, os algoritmos de replicação LCE e Prob( $p=0,7$ ) que replicam mais identidades, tendem a ter um desempenho pior em redes da névoa de menor capacidade se comparados com o LCD e Prob( $p=0,3$ ). A medida em que se é oferecido uma maior capacidade de armazenagem, os algoritmos LCE e Prob( $p=0,7$ ) tendem a se igualar e posteriormente ter um desempenho maior que o LCD e Prob( $p=0,3$ ). De uma forma geral, observa-se uma redução de até 90% do número de requisições para nuvem quando a política de substituição LRU é utilizada em conjunto com o algoritmo de replicação Prob( $p=0,7$ ). Esses resultados fornecem indícios de que uma replicação rápida das identidades, quando se possui uma grande capacidade de armazenagem na névoa, podem reduzir consideravelmente o número de requisições para a Nuvem.

Ainda com relação ao número de requisições, traçando um paralelo entre as Figuras 3 e 4 observa-se que um alto número de réplicas não é sinônimo de eficiência. Em redes da névoa de capacidade entre 1% e 5%, o Prob( $p=0,7$ ) reduz o número de identidades e o número de requisições para a nuvem em relação ao LCE. O Prob( $p=0,3$ ) por sua vez, apresenta número similares tanto no número de identidades quanto ao número de requisições para a nuvem similares ao LCD. Em redes da névoa de capacidade de 10%, o Prob( $p=0,3$ ) ofereceu um menor número de identidades e desempenho superior aos de-



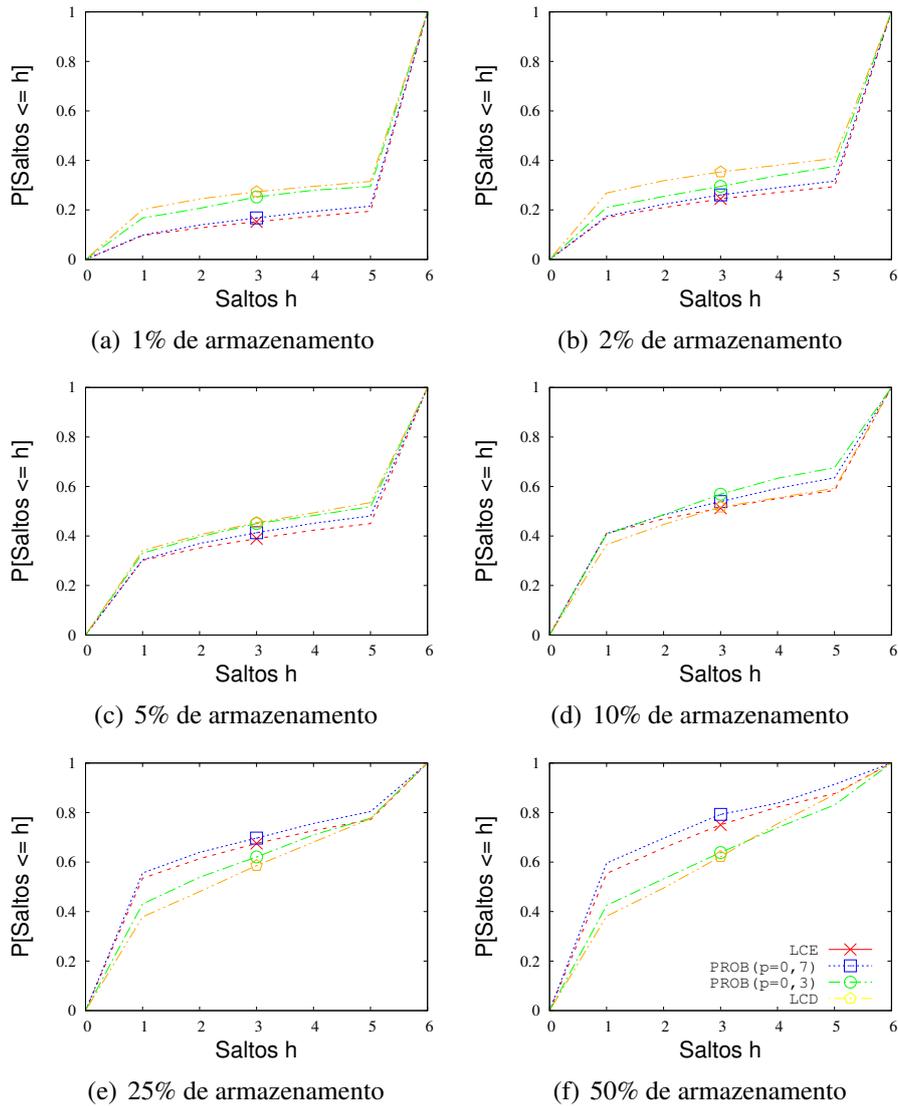
**Figura 4. Número de réplicas em relação a política de substituição de identidades**

mais. Por fim, em redes da névoa com capacidades superior a 25%, ocorre um impasse entre o número de réplicas e requisições para a nuvem. Enquanto o Prob( $p=0,7$ ) oferece um menor número de requisições para a nuvem, esse algoritmo requer mais identidades. O Prob( $p=0,3$ ) por sua vez, oferece cerca 5% a mais de requisições para a nuvem em troca de uma redução de 40% do número de réplicas em relação ao Prob( $p=0,7$ ).

### 6.3. Número de Saltos

A seguir, avaliamos o desempenho dos sistemas em relação ao número de saltos necessários para um dispositivo encontrar sua identidade. Essa análise permite verificar o desempenho dos sistemas em relação à posição das identidades na névoa. Diferente dos anteriores, para esses resultados fixa-se a política de replicação LRU, uma vez que essa apresentou um melhor desempenho em ambos casos anteriores.

A Figura 5 apresenta a função de distribuição acumulada do número de saltos necessários para encontrar as identidades. Note que o comportamento das políticas de replicação é muito variável em relação ao tamanho de armazenamento da rede da névoa. Novamente observa-se que um alto número de réplicas não é sinônimo de eficiência. Observe na Figura 5 (a) e (b) que o número médio de saltos necessários para encontrar a identidade é aproximadamente 15% menor nos algoritmos que replicam menos identidades LCD e Prob( $p=0,3$ ). No entanto, como observado na Figura 5 (c) e (d), a medida que se oferece mais recursos de armazenamento para a névoa, o número médio de saltos tende a se igualar em todos algoritmos. Em névoas com uma farta capacidade de armazenamento é possível notar uma inversão do comportamento dos algoritmos de replicação de identidade. Note que, os melhores algoritmos em névoas com menor capacidades (LCD e Prob( $p=0,3$ )), na Figura 5 (e) e (f), apresentou um número médio de saltos inferior ao LCE



**Figura 5. Saltos necessários para encontrar as identidades**

e Prob( $p=0,7$ ). Neste cenário, o Prob( $p=0,7$ ) apresentou queda de 5% do número médio de saltos necessários para se encontrar a identidade se comparado com o LCE, e 25% se comparado com o LCD. Contudo, quando consideramos o número de réplicas apresentado na Figura 3, o Prob( $p=0,7$ ) gerou um número inferior de identidades em relação ao LCE e 50% mais de identidades em relação ao LCD. Logo, nossos resultados mostram que a eficiência dos algoritmos de replicação de identidades nas névoas está relacionada a capacidade de armazenamento da névoa e a diversidade nas caches. Percebe-se que algoritmos baseados em um alto número de replicação possuem um desempenho melhor a medida em que se aumenta recursos, no entanto o Prob( $p=0,7$ ) ofereceu um menor número de identidades com maior diversidade nas redes da névoa se comparado com o LCE. Os algoritmos que replicam menos identidades possuem melhor desempenho quando utilizados em névoas com recursos limitados, em que, tanto o Prob( $p=0,3$ ) quanto o LCD geram números similares de réplicas das em todos cenários.

## 7. Conclusões

Neste artigo apresentamos uma investigação sobre o sistemas de distribuição de identidades em névoas. A eficiência da névoa para o armazenamento das identidades foi comprovada. Em cenários com alta capacidade de armazenamento, a distribuição de identidades foi capaz de reduzir em 90% o número de requisições para a nuvem. No entanto, nossa principal conclusão é que sistemas que se apoiam em algoritmos de replicação probabilísticos possuem um desempenho melhor que os demais. Ainda que exista um impasse entre a melhor probabilidade de replicação das identidades, de maneira geral, esses algoritmos possuem valores equilibrados entre o número de réplicas e as requisições encaminhadas para as nuvens. Em relação às políticas de replicação, apesar do impacto no sistema ser relativamente menor comparado com o algoritmo de distribuição, a política LRU foi capaz de reduzir em até 10% o número de requisições para a nuvem, implicando em uma maior eficiência da névoa quando utilizada com LRU. Em relação aos algoritmos de distribuição probabilísticos, quando utilizados com probabilidade de 70% em névoas com alta capacidade de armazenamento, esse algoritmo foi capaz de oferecer, com qualidade similar em relação ao número de saltos, um número 30% menor de identidades replicadas se comparado com o LCE. Assim, observamos que existe uma correlação entre o algoritmo de distribuição de identidades e a capacidade de armazenamento da névoa. Ainda, o algoritmo probabilístico, quando ajustado propriamente em relação ao tamanho do armazenamento da névoa, foi capaz de requisitar e replicar menos identidades para a nuvem. Quanto ao comportamento dos algoritmos em relação à distancia média de encontro das identidades, essa métrica tende a ter relação com o aumento da capacidade de armazenamento das névoas. Desta forma, identificamos a eficiência do algoritmo de distribuição de identidades probabilístico no contexto de névoas. Trabalhos futuros incluem investigações de outros algoritmos de replicação e substituição de identidades em névoa.

## Agradecimentos

Os autores agradecem o apoio da UFPR, CAPES, CNPQ, RNP e *National Science Foundation* (NSF). Este trabalho contou com auxílio financeiro do projeto de cooperação EUA-Brasil HealthSense, sob o termo de cooperação #99/2017 RNP/FUNPAR/UFPR.

## Referências

- Abbas, N., Zhang, Y., Taherkordi, A., and Skeie, T. (2018). Mobile edge computing: A survey. *IEEE Internet of Things Journal*, 5(1):450–465.
- Adamic, L. A. (2000). Zipf, power-laws, and pareto-a ranking tutorial. *Xerox Palo Alto Research Center*, <http://ginger.hpl.hp.com/shl/papers/ranking/ranking.html>.
- Bernardini, C., Silverston, T., and Olivier, F. (2013). Mpc: Popularity-based caching strategy for content centric networks. In *IEEE International Conference on Communications (ICC)*, pages 3619–3623. IEEE.
- Bonomi, F., Milito, R., Zhu, J., and Addepalli, S. (2012). Fog computing and its role in the Internet of Things. In *Proceedings of the first edition of the MCC workshop on Mobile cloud computing*, pages 13–16. ACM.
- Chai, W. K., He, D., Psaras, I., and Pavlou, G. (2012). Cache “less for more” in information-centric networks. In *International Conference on Research in Networking*, pages 27–40. Springer.

- Cisco (2016). Global mobile data traffic forecast update, 2015–2020 white paper. *link: <http://goo.gl/y1TuVx>*.
- Dsouza, C., Ahn, G.-J., and Taguinod, M. (2014). Policy-driven security management for fog computing: Preliminary framework and a case study. In *IEEE International Conference on Information Reuse and Integration (IRI)*, pages 16–23. IEEE.
- Gomez-Cárdenas, A., Masip-Bruin, X., Marin-Tordera, E., Kahvazadeh, S., and Garcia, J. (2018). A resource identity management strategy for combined fog-to-cloud systems. In *IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pages 1–6. IEEE.
- Habiba, U., Masood, R., Shibli, M. A., and Niazi, M. A. (2014). Cloud identity management security issues & solutions: a taxonomy. *Complex Adaptive Systems Modeling*, 2(1):1–37.
- Jacobson, V., Smetters, D. K., Thornton, J. D., Plass, M. F., Briggs, N. H., and Braynard, R. L. (2009). Networking named content. In *International Conference on Emerging Networking Experiments and Technologies*, pages 1–12. ACM.
- Laoutaris, N., Che, H., and Stavrakakis, I. (2006). The LCD interconnection of lru caches and its analysis. *Performance Evaluation*, 63(7):609–634.
- Neto, A. L. M., Souza, A. L. F., Cunha, I., Nogueira, M., Nunes, I. O., Cotta, L., Gentile, N., Loureiro, A. A. F., Aranha, D. F., Patil, H. K., and Oliveira, L. B. (2016). Aot: Authentication and access control for the entire iot device life-cycle. In *ACM Conference on Embedded Network Sensor Systems*, pages 1–15.
- Paul, S., Pan, J., and Jain, R. (2011). Architectures for the future networks and the next generation internet: A survey. *Computer Communications*, 34(1):2–42.
- Pсарas, I., Chai, W. K., and Pavlou, G. (2012). Probabilistic in-network caching for information-centric networks. In *Proceedings of the second edition of the ICN workshop on Information-centric networking*, pages 55–60. ACM.
- Roman, R., Lopez, J., and Mambo, M. (2018). Mobile edge computing, fog et al.: A survey and analysis of security threats and challenges. *Future Generation Computer Systems*, 78:680–698.
- Rossi, D. and Rossini, G. (2011). Caching performance of content centric networks under multi-path routing (and more). *Relatório técnico, Telecom ParisTech*, pages 1–6.
- Sabena, F., Dehghantanha, A., and Seddon, A. P. (2010). A review of vulnerabilities in identity management using biometrics. In *International Conference on Future Networks*, pages 42–49. IEEE.
- Torres, J., Nogueira, M., and Pujolle, G. (2013). A survey on identity management for the future network. *IEEE Communications Surveys & Tutorials*, 15(2):787–802.
- Trnka, M. and Cerny, T. (2016). Identity management of devices in internet of things environment. In *International Conference on IT Convergence and Security (ICITCS)*, pages 1–4. IEEE.
- Wu, H., Li, J., Pan, T., and Liu, B. (2013). A novel caching scheme for the backbone of named data networking. In *IEEE International Conference on Communications (ICC)*, pages 3634–3638. IEEE.