



Proposta para Grupo de Trabalho 2020

GT- RLProviDe-MI: feRramenta inteLigente, ágil e escalável para Provisionamento de recursos em reDes com Múltiplos Inquilinos

Dianne Scherly Varela de Medeiros

03 de Maio de 2020

1. Título

GT- RLProviDe-MI: feRramenta inteLigente, ágil e escalável para Provisionamento de recursos em reDes com Múltiplos Inquilinos

2. Coordenador Acadêmico

Nome: Dianne Scherly Varela de Medeiros

Instituição: Universidade Federal Fluminense

Currículo: Lattes - <http://lattes.cnpq.br/8119805151400395>

ResearchGate - [https://www.researchgate.net/profile/Dianne\\_Medeiros](https://www.researchgate.net/profile/Dianne_Medeiros)

Contato: E-mail - [diannescherly@id.uff.br](mailto:diannescherly@id.uff.br)

3. Assistente de Inovação

Nome: Pedro Silveira Pisa  
Instituição: Solvimm  
Universidade Federal Fluminense  
Currículo: LinkedIn - <https://www.linkedin.com/in/pedropisa/>  
Lattes - <http://lattes.cnpq.br/1898925829015070>  
Contato: E-mail: [pedro.pisa@solvimm.com](mailto:pedro.pisa@solvimm.com)

#### 4. Tema(s)

Serviços Inteligentes ligados ao Monitoramento e Medições da Rede: Monitoramento de desempenho do plano de dados da rede; Orquestração, automação e virtualização em redes

#### 5. Resumo

A alocação eficiente de banda em nuvens é desafiadora devido ao compartilhamento dos recursos entre inquilinos. Isso pode provocar ociosidade caso os inquilinos utilizem apenas a banda contratada. A alocação dinâmica de recursos sob demanda permite otimizar o uso da nuvem. Assim, as nuvens se beneficiam do uso de ferramentas baseadas em aprendizado de máquina que promovam respostas adaptáveis a ambientes variantes no tempo. Ademais, o uso de lógica difusa aumenta o nível de automação. Este projeto propõe uma ferramenta baseada em aprendizado por reforço e lógica difusa, capaz de gerenciar o acesso dos inquilinos aos recursos de rede da nuvem de forma ágil. O objetivo é maximizar a receita do provedor e reduzir a ociosidade da nuvem enquanto garante o nível de serviço contratado por inquilinos prioritários.

#### 6. Abstract

Efficient traffic allocation in clouds is challenging due to the resource sharing among tenants. This can cause idleness if tenants use only their hired bandwidth. On-demand based dynamic resource allocation allows to optimize the cloud usage. Hence, clouds benefit from the use of tools based on machine learning techniques that promote adaptive responses to time-varying environments. Besides, the use of fuzzy logic increases the automation level. This project proposes a tool based on reinforcement learning and fuzzy logic, capable of managing tenants' access to cloud resources in an agile fashion. The goal is to maximize the provider's profit and reduce cloud idleness while guaranteeing the service level agreements of priority tenants.

#### 7. Parcerias e respectivas contrapartidas

O Laboratório MídiaCom está iniciando colaboração com a empresa Solvimm (<https://solvimm.com/>) através do seu sócio-fundador, Pedro Silveira Pisa, M.Sc., que atualmente é aluno de doutorado no Laboratório Mídiacom pelo Programa de PósGraduação em Engenharia Elétrica e de Telecomunicações da Universidade Federal Fluminense. A Solvimm é especializada em soluções de computação em nuvem, big data e segurança da informação, e realiza consultorias para o desenvolvimento de visão de negócios e produtos no mercado de desenvolvimento de software e computação em nuvem. A empresa também possui parceria com a Amazon Web Services (AWS), oferecendo software de computação em nuvem e big data. Como contrapartida, a Solvimm ajudará na prospecção de créditos educacionais para uso de serviços da AWS proporcional ao valor repassado para a empresa na forma de pagamento de pessoal. Ademais, a empresa parceira fornecerá consultoria na implantação da solução proposta.

O Laboratório MídiaCom também desenvolve parceria com o pesquisador Martin Andreoni, D.Sc., da empresa Samsung Electronics, líder global em tecnologia da informação. A parceria vem se firmando através de colaborações em diversas áreas da tecnologia da informação e comunicação, sendo o foco atual a pesquisa e desenvolvimento de soluções baseadas em

inteligência artificial, centradas em processamento de dados em fluxo em tempo real. A contrapartida do parceiro ocorrerá na forma de consultorias relacionadas ao desenvolvimento de soluções baseadas em inteligência artificial para alocação de recursos em redes com infraestrutura compartilhada.

#### 8. Descrição do problema e da solução proposta com destaque para as inovações

A computação em nuvem vem sendo cada vez mais utilizada para diferentes finalidades. Nesse cenário, coexistem diversas infraestruturas físicas e virtuais, cada uma com seus requisitos de processamento, memória e armazenamento. A alocação de recursos é fundamental para os provedores de infraestrutura como serviço (Infrastructure as a Service – IaaS), responsáveis pela gerência e manutenção dos centros de dados, assim como para os inquilinos (tenants), que alugam recursos de centros de dados para implantar seus serviços na nuvem. Há grande diversidade nas demandas dos inquilinos, cada uma com exigências de capacidade diferentes durante períodos distintos (Chen et al., 2019, Medeiros et al., 2019). O uso eficiente dos recursos para o provedor IaaS incorre na multiplexação dos diversos inquilinos sobre os recursos físicos e é significativo para sua receita. Isso porque a maior utilização de recursos indica que um número maior de requisições dos inquilinos estão sendo atendidas e, portanto, maior lucro é gerado. Para os inquilinos coexistentes, os recursos alocados determinam o desempenho dos serviços implantados na nuvem e incorrem em custos (Chen et al., 2014). Atender às demandas de processamento, armazenamento, memória e rede, fornecendo o desempenho requerido pelos inquilinos e maximizando o lucro dos provedores é um problema NP-difícil (Guo et al., 2017). Dessa forma, a natureza dinâmica do sistema e a existência de diferentes requisitos dificultam a alocação eficiente de recursos, que continua sendo um desafio em aberto.

Os provedores IaaS devem ser capazes de prover o desempenho mínimo requerido pelos inquilinos, mesmo quando a nuvem está próxima de sua carga máxima suportada (Fox et al., 2009, Carvalho et al., 2011). Paralelamente, recursos ociosos, provenientes de capacidade disponível e não alocada, representam perdas de receita caso inquilinos com requisitos já atendidos solicitem mais recursos em um determinado período e não sejam atendidos. Nesse contexto, surge o modelo de venda sob demanda (pay-as-you-go), no qual inquilinos que contratam uma capacidade menor podem excedê-la, caso haja disponibilidade por parte da infraestrutura da nuvem, desde que o inquilino pague uma tarifa adicional. Assim, são necessários mecanismos que permitam distribuir os recursos do conjunto de todos os inquilinos requisitando recursos da nuvem, de forma a atender o desempenho mínimo de cada inquilino e, simultaneamente, redistribuir a capacidade ociosa da nuvem, aumentando o lucro do provedor. Os provedores de infraestrutura oferecem a seus inquilinos recursos computacionais sob demanda para alocação de máquinas virtuais. Diferentemente de recursos de processamento e memória que são compartilhados localmente em cada servidor, o recurso de rede é compartilhado globalmente entre os inquilinos que compartilham enlaces na rede. Assim, o compartilhamento da banda alocada para cada inquilino depende das tarefas que estão alocadas no mesmo servidor, dos inquilinos que estão se comunicando com ele e dos inquilinos que passam pelo mesmo enlace de gargalo. Não há, portanto, garantias de desempenho para recursos de rede, o que acarreta imprevisibilidade do desempenho das aplicações, que pode gerar insatisfação por parte dos inquilinos e causar perda de receita para o provedor IaaS (Bari et al., 2012). Para cada inquilino, variações no desempenho da rede dificultam a previsão do provedor e limitam seu custo máximo. Sem isolamento de desempenho de rede, um inquilino malicioso pode aumentar seu uso de rede em detrimento de outros. Esse problema reside no compartilhamento da rede baseado no serviço de melhor esforço do IP. A alocação de banda, então, diferencia-se da alocação dos recursos de processamento e memória, incorrendo em maiores dificuldades e acarretando desafios.

É possível provisionar recursos em situações previsíveis com planejamento. A automatização da alocação dos recursos em situações não previsíveis (Dutreilh et al., 2011) requer que o provisionamento automático se adapte às variações da rede, levando em consideração a

heterogeneidade dos requisitos. As propriedades da alocação de banda em provedores IaaS em nuvens são as garantias determinísticas, a justiça, a utilização e a praticabilidade (Chen et al., 2014). Garantias determinísticas de largura de banda asseguram um desempenho previsível para cada inquilino, independentemente de qualquer comportamento dos inquilinos concorrentes. As garantias mínimas de largura de banda permitem aos inquilinos obter mais largura de banda além de suas garantias quando há largura de banda disponível. Isso permite que os inquilinos obtenham desempenho limitado em pior caso e, portanto, custos limitados. A justiça é capaz de fornecer proteção relativa aos inquilinos bem-comportados com a presença de mau comportamento. Assim, garante a alocação justa de banda mesmo em cenários de disputa. A utilização relaciona-se ao uso da banda disponível no centro de dados. Alta utilização indica que a largura de banda não deve ficar ociosa enquanto houver demandas insatisfeitas. Um serviço com alta demanda de rede pode usar toda a largura de banda disponível quando outros serviços estão inativos. Garantir a utilização implica melhor desempenho de inquilinos cujos serviços têm demanda de tráfego atendidas. Provedores IaaS também se beneficiam de alta utilização da largura de banda, com a capacidade de acomodar mais solicitações de inquilinos e, assim, gerar mais renda. Por fim, a praticabilidade visa garantir a aplicação simplificada dos procedimentos de alocação de banda. Em especial a praticabilidade relaciona-se com a simplicidade e a escalabilidade da alocação de banda na nuvem.

O objetivo geral deste projeto é promover a melhoria de desempenho de uma nuvem compartilhada por múltiplos inquilinos ao mesmo tempo em que melhora a receita do provedor de IaaS. Para tanto, os seguintes objetivos específicos são estabelecidos: (i) desenvolver agentes de aprendizado por reforço, responsáveis por repassar as ações tomadas para um controlador; (ii) desenvolver um controlador baseado em lógica difusa, cuja função é determinar a banda disponível para cada inquilino de acordo com as ações dos agentes individuais, e determinar os hiperparâmetros do modelo de aprendizado por reforço para otimizar o desempenho do algoritmo; e (iii) otimizar a receita do provedor IaaS sem provocar não conformidade com o nível de serviço contratado. Dessa forma, o Produto Mínimo Viável identificado neste projeto é a realização de um protótipo de uma ferramenta para alocação de recursos em um ambiente altamente dinâmico. A ideia principal da ferramenta é garantir o aprendizado automático da política que maximiza o lucro do provedor IaaS, ao mesmo tempo em que garante o desempenho mínimo requerido por cada inquilino. Assim, a proposta deste projeto visa assegurar garantias determinísticas de que o tráfego mínimo contratado por um inquilino será sempre atendido, mesmo na presença de outros inquilinos que utilizem mais recursos do que o contratado. O mecanismo proposto garante a justiça ao distribuir a banda ociosa proporcionalmente à prioridade dos inquilinos que disputam o uso da banda. A utilização é otimizada na proposta, pois um dos objetivos é garantir a alta utilização da banda da nuvem para atender o máximo de solicitações por banda possível. Por fim, a praticabilidade da proposta é ressaltada pela sua simplicidade e pelo baixo consumo de memória e processamento, implicando escalabilidade. O desenvolvimento do produto mínimo viável e seu consequente aprimoramento impulsionará ainda mais a visibilidade da empresa parceira, aumentando seu rol de produtos com o objetivo de se tornar referência no desenvolvimento e implantação de soluções completas e seguras em ambientes de computação em nuvem e big data. É importante destacar que a ferramenta proposta beneficia os Sistemas RNP como o Compute@RNP, melhorando seu desempenho e a satisfação dos usuários. Além disso, uma vez que a solução proposta é compatível com ambientes de redes definidas por software, e com a implementação do projeto Infraestrutura Definida por Software (IDS) que está em andamento na RNP, a própria Rede Ipê se beneficia da proposta deste projeto.

Diversos trabalhos existentes na literatura propõem soluções para a alocação dinâmica de recursos em nuvens. As soluções utilizam técnicas proativas e reativas. Exemplos dessas técnicas são as baseadas em políticas que impõem limites, em análise de séries temporais, na teoria de filas, na teoria de controle e no aprendizado por reforço. As soluções que implementam limitações através de regras estabelecidas em políticas constituem uma abordagem reativa aparentemente simples.

A técnica é popular e é utilizada em sistemas comerciais como Amazon EC2<sup>1</sup> e Cloud Management Platform (Flexera)<sup>2</sup>. No entanto, o processo de configuração de limites é feito por aplicação, tornando-se complexo, e requer conhecimento profundo dos padrões de carga na nuvem para que não ocorra instabilidade do sistema. A análise de séries temporais (Botran et al., 2014) é uma abordagem proativa capaz de determinar os padrões de repetição da carga de entrada no sistema com o objetivo de prever o comportamento futuro. A acurácia da previsão depende do tamanho da janela de dados históricos analisada e da carga de entrada. Geralmente se utiliza estatística linear baseada em modelos de autorregressão, como Box e Jenkins (Box e Jenkins, 1990), implementados em algoritmos como ARIMA e ARMAX. Porém, a série histórica pode seguir um padrão não linear, de forma que as características da carga de entrada podem não ser bem capturadas pelos modelos lineares.

A teoria de filas (Botran et al., 2014, Eldin et al. 2012) é usada tanto em abordagens proativas quanto reativas, embora seja destinada para sistemas estacionários. Em sistemas dinâmicos com cargas heterogêneas, o modelo de fila precisa, então, ser periodicamente recalculado. A teoria de controle (Patikirikorala e Colman, 2010) destina-se a soluções reativas, sendo estendida para abordagens proativas através do emprego de modelos preditivos. Frequentemente usada para automatizar o gerenciamento de sistemas de servidores web e de centros de dados ou clusters, a teoria de controle é apropriada para ambientes em que a variação de carga é lenta, falhando na captura em ambientes dinâmicos quando rajadas repentinas chegam ao sistema. O aprendizado por reforço, por sua vez, garante uma utilização

---

estável do sistema mesmo sob condições de carga altamente dinâmica (Heinze et al., 2014). Além disso, os sistemas baseados em aprendizado por reforço identificam o modelo de desempenho e a política de um determinado ambiente sem nenhum conhecimento anterior. Assim, o aprendizado por reforço se torna uma solução robusta e adaptativa para o problema de alocação automatizada de recursos. Ainda, as soluções baseadas em aprendizado por reforço têm a vantagem de serem agnósticas à aplicação. Logo, diferentes aplicações podem usar o mesmo algoritmo sem a necessidade de haver personalização de configurações ou parâmetros. Em geral, as aplicações baseadas em aprendizado por reforço sofrem com a convergência lenta, não escalabilidade e desempenho fraco durante o período de aprendizado. Portanto, é necessário desenvolver uma solução eficiente que apresente convergência em tempo ótimo para que o desempenho da solução seja efetivo em um ambiente dinâmico. Dessa forma, ainda existe um grande desafio científico para desenvolver aplicações eficientes baseadas em aprendizado por reforço em sistemas altamente dinâmicos que resultem em bom desempenho.

Na literatura científica, muitos trabalhos utilizam aprendizado por reforço para solucionar o problema da alocação. Em geral se utiliza múltiplos agentes para controlar a alocação (Russell e Zimdars, 2003). Em algumas soluções, os estados são modelados de acordo com a proporção de inquilinos com uma ou mais requisições na fila e a priorização é realizada de acordo com a distância do período de tarifação (Mera-Gómez et al., 2017). As soluções existentes costumam focar no uso do algoritmo Q-Learning, por apresentar melhor desempenho do que outros algoritmos de aprendizado por reforço, como o SARSA (State-Action-Reward-State-Action) (Habib e Khan, 2016). Outras soluções apontam para o uso de Deep Q-Learning, buscando balancear os custos e o desempenho da nuvem (Wang et al., 2017). Para reduzir o tempo de convergência do Q-Learning e de outros algoritmos de aprendizado por reforço, é possível utilizar paralelismo na solução, compartilhando as experiências de aprendizado entre os agentes (Barrett et al., 2013). Outra abordagem é a utilização de conhecimento sobre o domínio da aplicação, inicializando parâmetros do modelo com base nesse conhecimento. Algoritmos de aprendizado

---

<sup>1</sup> <https://aws.amazon.com/pt/ec2/>

<sup>2</sup> Flexera previamente conhecida como RightScale, <https://www.flexera.com/products/agility/cloudmanagement-platform.html>

por reforço que convergem em tempo polinomial como o Speedy Q-Learning (Azar et al., 2011) também podem ser usados.

A ferramenta proposta neste projeto promove a alocação ágil e dinâmica do tráfego de uma nuvem. Vislumbra-se a implementação de múltiplos agentes que executam a alocação dinâmica de banda na nuvem através de uma estratégia baseada no algoritmo Q-learning. Os múltiplos agentes são capazes de limitar a taxa de transmissão de cada inquilino, mantendo, no entanto, a adequação ao desempenho mínimo de cada um, garantindo, assim, os níveis de serviço contratados. Os agentes são independentes, uma vez que a decisão de outro agente não influencia nas ações executadas pelos demais. Todos os agentes compartilham as mesmas informações e estados, sendo que cada agente tem controle sobre a banda disponível para um único inquilino. Além disso, existe uma restrição comum a todos os agentes de que não é permitido que o tráfego da nuvem aumente além de um limite estabelecido. A decisão de cada agente sobre a limitação do tráfego é tomada de acordo com os interesses do inquilino ao qual está atribuído e da capacidade atual da nuvem. Para a implementação proposta, assume-se que o problema segue um processo de decisão markoviano (Markov Decision Process – MDP), no qual os agentes são independentes, já que a decisão dos outros agentes não tem influência na ação tomada, havendo somente a restrição de que o tráfego total direcionado à nuvem não pode ultrapassar um limite pressuposto. A proposta se diferencia das demais existentes ao utilizar múltiplas políticas para garantir o cumprimento dos níveis de serviço contratados, otimizar o desempenho de cada agente para melhorar o desempenho geral do método, utilizar apenas informações sobre consumo de processamento e banda, e implementar um controle com base em lógica difusa para determinar os hiperparâmetros do modelo. Além disso, o modelo implementado é capaz de chavear entre as políticas de forma automática com base no conhecimento adquirido ao longo do tempo. A proposta inova ao combinar aprendizado por reforço e lógica difusa para maximizar a receita de um provedor IaaS, garantindo ainda os níveis mínimos de serviço contratados, sendo aplicada à alocação de recursos de rede nos provedores. É válido ainda enfatizar que a ferramenta proposta é de código aberto, podendo ser modificada e melhorada pela RNP e por outros usuários que tenham interesse em utilizá-la. Destaca-se também que, no contexto dos Sistemas RNP e grupos de trabalho recentes que se preocupam com o provisionamento de recursos de computação em nuvem distribuída, como GTSOFTWARE4IoT, GT-FENDE e TVoD-BrainIT, a proposta tem como principal diferencial a preocupação com o uso dos recursos de rede, uma questão não abordada pelos grupos citados.

#### 9. Ambiente de validação da solução proposta e documentação dos aprendizados

A validação da proposta ocorre em duas etapas. Na primeira, desenvolve-se os códigos e testes de validação usando uma infraestrutura virtualizada na qual se implementa um ambiente SDN. A utilização da SDN, no entanto, não limita a abrangência da proposta. A infraestrutura utilizada na primeira etapa se encontra no Laboratório Multiuso de Ensino de Redes de Nova Geração (LabGen), associado ao Laboratório MídiaCom da Universidade Federal Fluminense. Os equipamentos requisitados neste projeto complementarão a infraestrutura existente no LabGen para a execução dos múltiplos agentes e do controlador que compõem a ferramenta proposta no projeto. Essa etapa é essencial para o aperfeiçoamento do modelo de aprendizado por reforço e a lógica difusa usada pelo controlador para executar a alocação dinâmica de banda. O controlador é responsável por controlar a banda dos inquilinos com base nas ações recebidas dos agentes e por definir a troca de política de operação dos agentes de forma automática, através da modificação dos hiperparâmetros do modelo de aprendizado.

Com a modelagem matemática aperfeiçoada e com o aprimoramento do código visando eficiência, a segunda etapa de validação é executada. Na segunda etapa, a proposta é validada com a implantação em uma plataforma de nuvem de mercado, como a Amazon AWS, para testes de escalabilidade e desempenho em cenários mais densos. Nessa etapa, a parceria com a Solvimm é essencial para a prospecção de créditos educacionais junto à plataforma Amazon AWS e para o amparo e consultoria na implantação da solução.

## 10. Cronograma de marcos

Atividades e Entregas		Mês												
		2020						2021						
		JUN	JUL	AGO	SET	OUT	NOV	DEZ	JAN	FEV	MAR	ABR	MAI	JUN
<b>A1</b>	Especificação de equipe e equipamentos													
	E1   RP0 + PT até 17/06/2020													
<b>A2</b>	Elaboração de relatórios mensais de atividades													
	E2   RA até o último dia útil de cada mês													
<b>A3</b>	Reunião inicial													
	E3   RP1 entre 03/08 e 14/08/2020													
<b>A4</b>	Webconferência de alinhamento sobre contrapartidas													
	E4   RP2 entre 17/08 e 31/08/2020													
<b>A5</b>	Elaboração do relatório de prospecção													
	E5   RT1 até 30/09/2020													
<b>A6</b>	Desenvolvimento da capacidade empreendedora													
	E6   RP3 entre 01/08 e 31/10/2020													
<b>A7</b>	Elaboração do relatório da visão de negócios e produto													
	E7   RP4 + RT2 até 30/10/2020													
<b>A8</b>	Desenvolvimento com mentorias													
	E8   SM entre 01/11/2020 e 31/07/2021													

(Continuação da tabela)

Atividades e Entregas		Mês												
		2020						2021						
		JUN	JUL	AGO	SET	OUT	NOV	DEZ	JAN	FEV	MAR	ABR	MAI	JUN
<b>A9</b>	Desenvolvimento de landing page e vídeo pitch													
	E9.1   Landing page até 03/11/2020													
	E9.2   Vídeo pitch até 03/11/2020													
<b>A10</b>	Webinar de apresentação para RNP													
	E10   Apresentação executiva entre 04/11 e 19/11/2020													
<b>A11</b>	Acompanhamento dos ciclos de desenvolvimento e validação do MVP													
	E11.1   Evidências da validação do modelo de negócios, mensalmente entre 01/11/2020 e 30/04/2021													
	E11.2   Evidências do desenvolvimento tecnológico, mensalmente entre 01/11/2020 e 30/04/2021													
<b>A12</b>	Desenvolvimento da versão preliminar do MVP													
	E12   Versão preliminar do MVP até 03/05/2021													
<b>A13</b>	Elaboração do whitepaper do MVP													
	E13   Whitepaper do MVP até 03/05/2021													
<b>A14</b>	Demonstração no Workshop RNP (WRNP)													
	E14   Demonstração do MVP entre 03/05 e 31/05/2021													
<b>A15</b>	Preparação da apresentação final do projeto													
	E15   Evento para apresentação do projeto finalizado em 01/06 e 30/06/2021													
<b>A16</b>	Disponibilização do código-fonte da versão final do projeto													
	E16   Código-fonte em 01/07/2021													
<b>A17</b>	Webconferência para planejamento e registro de software													
	E17   RP5 entre 01/06/2021 e 31/07/2021													

- RP Relatório de Planejamento
- PT Plano de Trabalho
- RA Relatório de Acompanhamento
- RT Relatório Técnico

## 11. Recursos financeiros

### 11.1.1. Equipamentos

Os equipamentos solicitados para o desenvolvimento do projeto deste GT estão listados na tabela a seguir. Os equipamentos seguem a configuração padrão descrita no Anexo 4 do Edital de GTs 2020 da RNP.

Descrição	Instituição de Destino	Qtd.	Valor em R\$ estimado
Desktop Modelo i7 (Core i7 - 16GB - SSD 256GB) + Kit teclado e mouse com fio Incluso: Garantia 3 anos ProSupport, troca de peças onsite	Universidade Federal Fluminense	2	7.205,28
Notebook 14" Modelo i7 (Core i7 - 8GB - SSD 256GB) Incluso: Garantia 3 anos Complete Care, com cobertura acidental e troca de peças onsite	Universidade Federal Fluminense	3	15075,72
Monitor 24"	Universidade Federal Fluminense	2	2099,58
Subtotal			24.380,58

## 12. Referências

- Bari, M. F., Boutaba, R., Esteves, R., Granville, L. Z., Podlesny, M., Rabbani, M. G., Zhang, Q. e Zhani, M. F. (2012). Data center network virtualization: A survey. *IEEE Communications Surveys & Tutorials*, 15(2):909–928.
- Barrett, E., Howley, E. e Duggan, J. (2013). Applying reinforcement learning towards automating resource allocation and application scalability in the cloud. *Concurrency and Computation: Practice and Experience*, 25(12):1656–1674.
- Botran TL, Alonso JM, Lozano JA (2014) A review of Auto-scaling Techniques for Elastic Applications in Cloud Environments. *Journal of Grid Computing* 12(4):559–592.
- Box G, Jenkins G (1990) *Time Series Analysis, Forecasting and Control*. Holden-Day, Incorporated, San Francisco
- Carvalho, H. E. T., Fernandes, N. C. e Duarte, O. C. M. B. (2011). Um controlador robusto de acordos de nível de serviço para redes virtuais baseado em lógica nebulosa. Em: *Anais do XXIX Simpósio Brasileiro de Redes de Computadores SBRC'11*, Campo Grande, MS.
- Chen, L., Li, B. e Li, B. (2014). Allocating bandwidth in datacenter networks: A survey. *Journal of Computer Science and Technology*, 29(5):910–917.
- Chen, Z., Hu, J. e Min, G. (2019). Learning-based resource allocation in cloud data center using advantage actor-critic. Em: *Proceedings of the International Conference on Communications (ICC)*, p. 1–6. IEEE.
- Dutreilh, X., Kirgizov, S., Melekhova, O., Malenfant, J., Rivierre, N. e Truck, I. (2011). Using reinforcement learning for autonomic resource allocation in clouds: towards a fully automated workflow. Em: *Proceedings of the Seventh International Conference on Autonomic and Autonomous Systems (ICAS)*, p. 67–74.
- Eldin, A. A., Tordsson, J., Elmroth, E. (2012) An adaptive hybrid elasticity controller for cloud infrastructures, Em: *Network Operations and Management Symposium (NOMS)*, p. 204–212.
- Fox, A., Griffith, R., Joseph, A., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A. e Stoica, I. (2009). Above the clouds: A Berkeley view of cloud computing. Dept. Electrical Eng. and Comput. Sciences, University of California, Berkeley, Rep. UCB/EECS, 28(13):2009.

- Guo, J., Song, Z., Cui, Y., Liu, Z. e Ji, Y. (2017). Energy-efficient resource allocation for multi-user mobile edge computing. Em: Proceedings of the Global Communications Conference (GLOBECOM), p. 1–7.
- Habib, A. e Khan, M. I. (2016). Reinforcement learning based autonomic virtual machine management in clouds. Em: Proceedings of the International Conference on Informatics, Electronics and Vision (ICIEV), p. 1083–1088.
- Heinze, M., Pappalardo, V., Jerzak, Z., Fetzer, C. (2014). Auto-scaling Techniques for Elastic Data Stream Processing. Em: Proceedings of the ACM International Conference on Distributed Event-Based Systems (DEBS), p. 318–321
- Medeiros, D. S. V., Cunha Neto, H. N., Andreoni Lopez, M., Magalhães, L. C. S., Silva, E. F., Vieira, A. B., Fernandes, N. C. e Mattos, D. M. F. (2019). Análise de dados em redes sem fio de grande porte: Processamento em fluxo em tempo real, tendências e desafios. Em: Minicursos do Simpósio Brasileiro de Redes de Computadores (SBRC), 2019:142–195.
- Mera-Gómez, C., Ramírez, F., Bahsoon, R. e Buyya, R. (2017). A debt-aware learning approach for resource adaptations in cloud elasticity management. Em: Proceedings of the International Conference on Service-Oriented Computing, p. 367–382.
- Azar, M. G., Munos, R., Ghavamzadeh, M. e Kappen, H. J. (2011). Speedy Q-learning. Em: Proceedings of the International Conference on Neural Information Processing Systems (NIPS), p. 2411–2419.
- Patikirikorala T, Colman A. (2010) Feedback controllers in the cloud. Em: Proceedings of Cloud Workshop, at Asia Pacific Software Engineering Conference (APSEC 2010), p. 19-24 .
- Russell, S. J. e Zimdars, A. (2003). Q-decomposition for reinforcement learning agents. Em: Proceedings of the International Conference on Machine Learning (ICML), p. 656–663.
- Wang, Z., Gwon, C., Oates, T. e Iezzi, A. (2017). Automated cloud provisioning on AWS using deep reinforcement learning. CoRR, abs/1709.04305, p.1–8.