

Revista Brasileira de Computação Aplicada, Abril, 2021

DOI: 10.5335/rbca.v13i1.10827 Vol. 13, N° 1, pp. 11–26

Homepage: seer.upf.br/index.php/rbca/index

ARTIGO ORIGINAL

Avaliação de Desempenho, Consumo de Energia e Custo para Ambientes Baseados em Contêineres e Máquinas Virtuais

Evaluation of Performance, Energy Consumption and Cost for Environments Based on Containers and Virtual Machines

Cleyton Ferreira Gonçalves ^{[0,1}, Ermeson Andrade ^{[0,1}, Gustavo Callou ^{[0,1}, Bruno Nogueira ^{10,2}

¹Departamento de Computação, Universidade Federal Rural de Pernambuco, Recife, Brazil, ²Instituto de Computação, Universidade Federal de Alagoas, Maceió, Brazil

*cleyton.goncalves@ufrpe.br; ermeson.andrade@ufrpe.br; gustavo.callou@ufrpe.br; bruno@ic.ufal.br

Recebido: 07/04/2020. Revisado: 31/08/2020. Aceito: 09/11/2020.

Resumo

A virtualização oferece melhorias significativas na utilização do recursos dos data centers, reduzindo o desperdício de energia em razão da redução da subutilização de máquinas físicas em operação. No entanto, o fator hypervisor e a (VM) acrescentam uma sobrecarga desnecessária no desempenho geral do sistema operacional hospedeiro, visto que cada instância baseada em VM possui seu próprio sistema operacional. Por outro lado, os contêineres têm se tornado uma alternativa leve e flexível aos ambientes cujas características mandatórias são agilidade e consistência nos processos de migração e replicação de aplicações. Diferentemente das tecnologias tradicionais de virtualização ou paravirtualização, os contêineres não exigem uma camada de emulação ou um hypervisor para serem executados, em vez disso, eles utilizam uma interface no nível do sistema operacional. Essa característica os torna uma tecnologia enxuta, visto que exigem uma sobrecarga limitada e permite que maiores densidades de instâncias sejam executadas no computador hospedeiro. Este artigo observou a utilização de CPUs e os custos da energia consumida por três ambientes experimentais baseados em contêineres Docker e máquinas virtuais do Kernel-based Virtual Machine (KVM). Os resultados demostraram que os contêineres apresentaram um melhor equilíbrio de aproximadamente 1,82% na utilização na CPU, contra 28,89% da VMs. Além disso, os resultados do consumo de energia e do custo por killowatt-hora dos contêineres foram significativamente

Palavras-Chave: Virtualização; Contêineres; Máquina Virtual; Avaliação de Desempenho; Consumo de Energia.

Abstract

Virtualization has provided significant improvements in data center resource utilization, since reduces energy waste due to the reduced underutilization of physical machines in operation. However, the hypervisor factor and the virtual machine (VM) add unnecessary overhead to the overall performance of the host operating system, because each VMbased instance has its own guest operating system. On the other hand, containers have become a lightweight and flexible alternative to environments whose mandatory characteristics are agility and consistency in migration processes and application replication. Unlike traditional virtualization or paravirtualization technologies, containers do not require an emulation layer or hypervisor to run, instead they use an operating system-level interface. This feature makes them a lean technology as they require limited overhead, allowing higher instance densities to run on the host computer. This paper aimed to analyze CPU utilization and energy consumption costs in three experimental environments based on Docker containers and Kernel-based Virtual Machine (KVM). The results showed that the containers had a better balance of approximately 1.82% in CPU utilization, compared to 28.89% of VMs. Besides, the energy consumption and cost per killowatt hour of containers were significantly lower.

Keywords: Virtualization; Containers; Virtual Machine; Performance Evaluation; Energy Consumption.

1 Introdução

As infraestruturas de *data centers* são ambientes que abrigam vários sistemas computacionais, como servidores, equipamentos de processamento, armazenamento de dados, dispositivos de rede, cabeamento estruturado, alimentação de energia e sistema refrigeração. Os provedores de tecnologia da informação e comunicação (TIC) do mercado (ex.: *Google, Facebook, Amazon, Netflix, Apple*, etc.) utilizam várias salas e edificações para construir *data centers* com capacidade de processamento e transmissão de dados compatível ao aumento da demanda de serviços em escala global. Alguns estudos apontaram que o aumento na demanda por tráfego na internet e dispositivos móveis está resultando no crescimento do consumo de energia e emissão de gases de efeito estufa por parte dos *data centers* (Jones, 2018).

Em Deka (2014), o autor projeta que o crescimento na demanda de conteúdo digital na internet atinja cerca de 35 ZBytes do tráfego IP até 2021. Uma pesquisa da Cisco (2019) estimou que as demandas no consumo dos serviços de streaming de vídeos representarão 82% de todo tráfego da internet até 2022. Tal projeção indica que os serviços de streaming elevarão os níveis de utilização dos recursos e consumo de energia dos data centers. A pesquisa de Belkhir and Elmeligi (2018) estimou que os equipamentos de TIC representam até 8% do consumo de energia global, e projeta que os mesmos atingirão cerca de 14% em 2021. O estudo de Jones (2018) apontou que os data centers atualmente consomem cerca de 200 terawatt-hora (tWh) de eletricidade por ano. Isso representa 1% da demanda global de eletricidade, superando o consumo de energia de alguns países.

O estudo de Delforge and Whitney (2014) projetou que o consumo de energia dos data centers dos EUA aumentará aproximadamente para 140 bilhões de quilowatts-hora por ano. Em termos monetários, esses valores custarão às empresas americanas cerca e US\$ 13 bilhões. Isso representa uma produção de quase 150 milhões de toneladas métricas de emissões de carbono por ano, o que equivale à produção anual e poluição de 50 usinas a carvão. De acordo com Jones (2018), a pegada de carbono global da indústria de TIC representa mais de 2% das emissões globais de carbono. Isso é comparável às emissões de combustíveis da indústria da aviação. Em até 2025, Andrae and Edler (2015) estimou um cenário onde a área de TIC será responsável por cerca de 8% da pegada de carbono global.

O advento da virtualização do hardware ofereceu melhorias significativas na utilização dos recursos dos *data centers*, eliminando o desperdício de energia em razão da redução da subutilização das máquinas físicas em operação (Bhimani et al., 2017). As máquinas virtuais (VM) são representações virtuais de computadores convencionais hospedados conforme disponibilidade dos recursos da máquina física. O *hypervisor* é uma camada de software intermediária que redireciona as instruções dos sistemas operacionais convidados (*guest OS*) de cada VM para o hardware da máquina física.

A virtualização otimizou a utilização dos recursos e reduziu a ociosidade das máquinas físicas. No entanto, o fator *hypervisor* e a máquina virtual acrescenta uma sobrecarga desnecessária no desempenho geral do sistema

operacional hospedeiro, visto que cada instância baseada em VM possui seu próprio sistema operacional (Tadesse et al., 2017). Por outro lado, a tecnologia contêiner preconiza baixa utilização dos recursos de hardware, visto que dispensa o uso do *hypervisor*, máquinas virtuais e sistemas operacionais convidados (Li et al., 2018), isto é, os contêineres são carregados no nível do sistema operacional (SO) da máquina física. As arquiteturas orientadas a serviços ou microsserviços utilizam os contêineres por serem compactos e leves. Tais características torna-os mais velozes e flexíveis na distribuição das aplicações e minimiza o custo das operações dos *data centers* (Bhimani et al., 2017).

Embora existam várias pesquisas disponíveis na literatura que analisam a performance dos contêineres e VMs, mas poucas exploram cenários focados no consumo de energia das operações baseadas nessas duas abordagens. Considerando que a eficiência energética é uma questão aberta no contexto dos data centers, este trabalho almeja realizar uma avaliação de desempenho e de consumo de energia em cenários baseados nas duas tecnologias. Os experimentos iniciais observarão a utilização da CPU e o consumo de energia da infraestrutura computacional proposta. As cargas de trabalhos serão variadas em três cenários compostos por contêineres e VMs de modo que a utilização dos recursos e a demanda da potência seja observada por amostras periódicas. O consumo de energia será coletado na unidade de potência Watt. No entanto, o consumo de energia oriundo dos experimentos serão expressos em kilowatt-hora, visto que essa unidade de potência é comumente aplicada nas cobranças de consumo de energia das concessionárias nacionais. Sendo assim, para facilitar entendimento da leitura, o fator custo será calculado a partir dos resultados obtidos nas medições das potências com base na tarifa média nacional do kilowatthora. O custo financeiro da demanda de potência média será apresentado através de projeções mensais e anuais. O ambiente experimental utilizará os contêineres Docker e as máquinas virtuais do hypervisor KVM. Os resultados das medições serão observados por meio de três cenários baseados nas duas tecnologias, onde cada cenário contemplará dois escopos compostos por quantidades de instâncias equivalentes.

O restante deste artigo está organizado da seguinte forma. A seção II descreve os fundamentos da virtualização e a conteinerização. A Seção III descreve os trabalhos relacionados à avaliação de desempenho e o consumo de energia dos ambientes baseados em contêineres e VMs. A Seção IV apresenta a arquitetura experimental, a configuração do ambiente de *testbed*, os cenários experimentais, os *benchmarks* e as ferramentas de medição. A Seção V apresenta uma discussão dos resultados. Finalmente, a Seção VI apresenta as conclusões e os trabalhos futuros.

2 Fundamentação Teórica

Esta seção introduz sucintamente os conceitos empregados na virtualização de máquinas virtuais e na conteinerização, bem como os softwares que aplicam as duas tecnologias nos contextos dos *data centers*. Na virtualização, o KVM (*Kernel Virtual Machine*) é uma ferramenta de software que implementa um conceito denominado *hypervisor*, o qual cria máquinas virtuais para carregar instâncias

de sistema operacionais independentes. O *Docker* é uma ferramenta de software que implementa um conceito denominado de conteinerização, a qual preconiza baixo consumo de recursos. Umas das características essenciais da conteinerização, é que as instâncias baseadas em contêineres carregam réplicas reduzidas de sistemas operacionais contendo apenas as bibliotecas necessárias à operação do serviço ou aplicação, uma vez que, o *kernel* do sistema operacional do computador físico (ou hospedeiro) é compartilhado com os contêineres.

2.1 Virtualização

A virtualização é uma tecnologia baseada em software que usa o compartilhamento do hardware da máquina física para reproduzir o comportamento de múltiplas abstrações de máquinas virtuais independentes e isoladas entre si (redhat, 2019). Em linha gerais, a virtualização funciona no sentido de otimizar a utilização do hardware e minimizar a ociosidade dos recursos computacionais. Sendo assim, é possível utilizar a capacidade total de uma máquina física, distribuindo partes menores dos recursos para múltiplas instâncias ou ambientes, como a virtualização de aplicativos, de computadores convencionais, de servidores de aplicação, de serviços de armazenamento, de serviços e funções de rede (VMware, 2019). Com virtualização, as empresas tendem a investir menos na aquisição de ambientes computacionais completos, em virtude das alocações dos recursos computacionais serem mais aprimoradas e econômicas.

O ambiente virtualizado é caracterizado pela agilidade e a flexibilidade nas entregas. Tais características aumentam a eficiência no aprovisionamento dos serviços ou aplicações. Além do disso, o gerenciamento do data center é simplificado através do aumento na automação das tarefas. Esse aspecto oferece mais mobilidade nas cargas de trabalho e, consequentemente, aumenta o desempenho e a disponibilidade dos serviços ou aplicações (Chirammal et al., 2016). De acordo com a VMware (2019), a virtualização oferece os seguintes benefícios: (a) redução dos custos operacionais e capital; (b) redução ou eliminação do tempo de inatividade; (c) aumento de produtividade, eficiência, agilidade e capacidade de resposta da TI; (d) rapidez no aprovisionamento dos aplicativos e recursos; (e) melhor continuidade dos negócios; (f) gerenciamento simplificado das operações dos data centers; (g) aumento na disponibilidade dos serviços e aplicações dos data centers.

A Fig. 1 demostra em termos práticos um cenário hipotético cuja infraestrutura é composta por três servidores físicos, sendo cada um deles com finalidades específicas. O primeiro é um servidor de *e-mail*, o segundo é um servidor *web* e o terceiro executa aplicações legadas internas. As demandas encaminhadas aos servidores de *e-mail*, *web*, legado utilizam cerca de 30% da capacidade dos recursos e 70% de ociosidade nos computadores físicos. Todavia, os valores elevados de ociosidade resultam em custos desnecessários que prejudicam a manutenção da operação do provedor de serviço. Considerando que os serviços de *e-mail* e *web* juntos utilizam cerca de 60% da capacidade dos recursos, um computador suportaria duas instâncias virtualizadas de ambos serviços e ainda restaria cerca 40% para eventuais elasticidades nas duas aplicações (redhat,

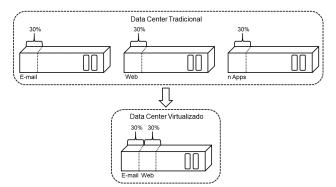


Figura 1: Processo de virtualização

2019).

Levando em conta que o computador físico da Fig. 1 adotou um esquema de virtualização que permite criar instâncias de computadores denominados máquinas virtuais, tanto o servidor de *e-mail* quanto o servidor *web* carregaram o seu próprio sistema operacional de forma independente e isolada. Sendo assim, as duas máquinas virtuais executaram simultaneamente diferentes sistemas operacionais compartilhando o hardware de um único computador físico (Chirammal et al., 2016). Dessa forma, a criação de múltiplas máquinas virtuais dentro de um único computador físico aprimora a escalabilidade, a elasticidade dos serviços e o balanceamento das cargas de trabalho.

De acordo com Ivanov (2017), a virtualização completa e a para-virtualização são as duas técnicas de virtualização amplamente adotadas nos ambientes de data centers. A adoção da implantação de uma delas varia conforme a necessidade da operação do negócio e a capacidade dos recursos. A virtualização completa (ou total) utiliza o hardware subjacente da máquina física para criar replicas completas que emulam dispositivos de hardware virtuais cujo comportamento se assemelha a um computador convencional que pode executar um sistema operacional convidado. Nessa técnica, os softwares, os sistemas operacionais, as aplicações e os serviços são executados diretamente no hardware emulado por um software de virtualização denominado monitor de máquina virtual (ou virtual machine monitor) que cria dispositivos virtuais com sistemas operacionais convidados. Sendo assim, o sistema hospedeiro ignora a existência da máquina virtual e opera como se funcionasse diretamente sobre o sistema operacional para o qual foi projetado. O resultado é um novo sistema virtual, no qual o sistema operacional convidado não sofre nenhuma modificação (Chirammal et al., 2016). Normalmente, a emulação desse ambiente é definido como hypervisor

A para-virtualização requer uma modificação nos sistemas operacionais convidados para que sejam realizadas as interações com o *hypervisor* e o processo de conversão das instruções sejam interpretadas por ele ou diretamente no hardware do computador hospedeiro. Nessa técnica, enquanto o sistema hospedeiro envia algumas instruções de baixo nível diretamente ao hardware hospedeiro, o *hypervisor* realiza a conversão das instruções de alto nível, intermediando a comunicação dos sistemas convidados e

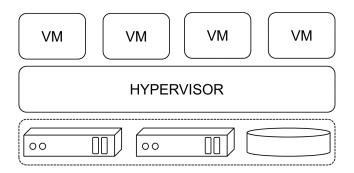


Figura 2: Um ambiente virtualizado por hypervisor

as camadas subjacentes. Os sistemas convidados da paravirtualização produzem menos sobrecargas no desempenho geral do sistema hospedeiro em relação aos sistemas convidados da virtualização completa (Chirammal et al., 2016).

A máquina virtual representa uma réplica isolada e eficiente de um computador real com sistema operacional, aplicativos e serviços (Cuadrado-Cordero et al., 2017). As VMs apresentam características como particionamento dos recursos do sistema entre múltiplas máquinas virtuais e a execução de diversos sistemas operacionais na máquina física, incluindo as características de isolamento de falhas, segurança no nível de hardware, preservação no desempenho do sistema, otimização na alocação e gerenciamento dos recursos. Os sistemas operacionais e as aplicações instanciados nas VMs são encapsulados em arquivos que podem ser movimentados ou replicados entre vários ambientes. Além dos mais, a independência do hardware garante um rápido aprovisionamento ou migração das aplicações contidas nas VMs para qualquer ambiente físico (Barik et al., 2016).

Conforme pode ser observado na Fig. 2, os hypervisors representam uma camada de software que dissocia as máquinas virtuais do computador hospedeiro e aloca os recursos de hardware dinamicamente ou estaticamente para cada instância virtualizada (Ivanov, 2017). Os hypervisors possuem implementações baseadas em softwares, firmwares e hardwares que permitem criar, executar e gerenciar as máquinas virtuais. Sendo assim, os hypervisors compartilham os recursos de hardware de uma máquina física entre várias máquinas virtuais e seus sistemas operacionais convidados. Os recursos de uma máquina física como CPUs, memórias, discos rígidos, interfaces de redes, periféricos, entre outros são virtualizados e replicados para vários sistemas operacionais carregados simultaneamente (Raho et al., 2015). As soluções mais conhecidas de hypervisors incluem o Kernel-based Virtual Machine KVM, o Xen, o QEMU, o VirtualBox, o VMware vSphere, o VMware Workstation e o Hyper-V.

Este artigo faz uso de um software de virtualização denominado de Máquina Virtual baseada em Núcleo (ou Kernel-based Virtual Machine). Na realidade, o KVM compila no kernel do Linux um módulo que implementa as funcionalidades de um hipervisor para máquinas virtuais. Esse módulo emula as funcionalidades de processamento e dispositivos periféricos, como: disco, rede, VGA, PCI, USB, portas seriais e paralelas entre outras, que representam

os componentes de hardwares virtuais reconhecidos pelos sistemas operacionais convidados instalados nas máquinas virtuais (Raho et al., 2015).

2.2 Conteinerização

A conteinerização é definida como um mecanismo que empacota arquivos, bibliotecas e dependências de aplicações por uma tecnologia denominada de contêineres, a qual abrange características de portabilidade, escalabilidade, configurabilidade, isolamento e economicidade na utilização dos recursos. Isso permite que as distribuições dos contêineres entre os ambientes de desenvolvimento e operação sejam conduzidas com máxima agilidade nas entregas, sem causar prejuízo a produção (Nickoloff and Kuenzli, 2019).

Os contêineres tornaram-se uma alternativa leve e flexível a ambientes cujas características mandatórias são agilidade e consistência nos processos de migração e replicação de aplicações entre os ambientes de desenvolvimento, teste, produção ou operação. Sendo assim, os contêineres oferecem processos mais práticos com o objetivo de maximizar a produtividade e o desempenho das operações (Barik et al., 2016).

Diferentemente das tecnologias tradicionais de virtualização ou paravirtualização, nos contêineres não exigem uma camada de emulação ou um hypervisor para serem executados, em vez disso, eles utilizam uma interface no nível do sistema operacional. Essa característica os torna uma tecnologia enxuta tendo em vista que exigem uma sobrecarga limitada para executá-los. Isso permite que maiores densidades de instâncias sejam executadas no computador hospedeiro (Turnbull, 2014). Além disso, os contêineres são carregados no espaço do usuário sobre o kernel do sistema operacional. Essa característica garante uma redução considerável na sobrecarga dos recursos e aumenta o desempenho das tarefas executados nos ambientes baseados em contêineres, visto que as instâncias são executadas diretamente no sistema operacional hospedeiro sem redirecionar as instruções para o hypervisor (Cuadrado-Cordero et al., 2017).

Os contêineres armazenam apenas os arquivos necessários à execução de um ou mais processos empacotados por uma imagem inicializada isoladamente no sistema operacional hospedeiro. A imagem inicializada no contêiner difere de um sistema operacional inicializado na VM, pois uma imagem no contêiner possui apenas as dependências e os binários para carregar os arquivos da aplicação, sem necessidade da intermediação do *hypervisor* que é uma ferramenta imprescindível na inicialização das VM e o seu respectivo sistema operacional convidado (Nickoloff and Kuenzli, 2019).

A Fig. 3 mostra um ambiente estruturado na virtualização e outro na conteinerização. Considerando que o mesmo hardware subjacente é compartilhado para ambas tecnologias, as camadas de softwares representam um ambiente virtualizado composto por um sistema operacional hospedeiro, um *hypervisor*, três máquinas virtuais e os seus respectivos sistemas operacionais convidados (*guest OS*). Já o ambiente conteinerizado é composto pelo sistema operacional hospedeiro, o mecanismo de distribuição (*docker enqine*), dois contêineres com seus respectivos processos

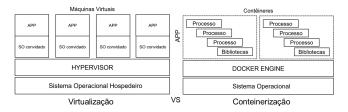


Figura 3: Virtualização vs. Conteinerização

e bibliotecas da aplicação.

O Docker é um mecanismo de código aberto que automatiza a implantação de aplicações baseadas em contêineres. Ele foi desenvolvido em 2008 pela equipe Docker e lançado sob a licença Apache 2.0. O Docker não é uma tecnologia de virtualização, trata-se de um serviço de gerenciamento de contêineres já incorporado no sistema operacional (Salah et al., 2017). A tecnologia Docker possui uma interface de gerenciamento que oferece diversos comandos para administrar as operações das imagens, um daemon do servidor, uma biblioteca de imagens compostas por contêineres predefinidos e comandos para armazenar, baixar e compartilhar imagens de servidores em repositórios públicos e privados (Nickoloff and Kuenzli, 2019). O Docker possibilita o empacotamento de uma ou mais aplicações ou ambientes completos dentro um ou mais contêineres, tornando-os portáveis para qualquer outro ambiente baseado na tecnologia. Isso reduz drasticamente o tempo de deploy das aplicações ou serviços. Outra facilidade do Docker é a criação de imagens a partir de arquivos contendo configurações predefinidas, denominados de Dockerfiles (Turnbull, 2014).

3 Trabalhos Relacionados

Esta seção apresenta os trabalhos relacionados a este artigo que aborda um estudo sobre a avaliação de desempenho, do consumo de energia e do custo de ambientes baseados em contêineres e VMs.

O estudo de Zheng et al. (2020) simulou no CloudSim4.0 um cluster baseado em contêineres utilizados nas avaliacões dos fatores consumo de energia e números de migrações. Os experimentos observaram alguns algoritmos de migrações dedicados às etapas de seleção da origem, seleção do contêiner e seleção de destino. A etapa de seleção da origem toma a decisão baseada na taxa utilização dos recursos conforme limites definidos estaticamente, dinamicamente através dos métodos Desvio Absoluto Médio (MAD) ou Intervalo Interquartil (IQR), modelos de predição como Regressão Linear (LR) ou Autoregressivo (AR) para evitar previamente desperdícios de recursos. Etapa de seleção dos contêineres toma a decisão baseada em três algoritmos: Utilização Máxima (MU) dos recursos para migrar menos contêineres para o nó, Correlação Máxima (MC) entre as taxas de utilização dos contêineres e nós, migração por Seleção Aleatória (RS). Por último, a etapa de seleção do destino observou o consumo de energia, a violação de SLA e os números de migrações de acordo com cinco algoritmos de agendamento de carga, a saber: Random Selection (RS), Minimum Energy Consumption (ME), Load Balancing (LB), Load Concentration (LC) e Resource

Load Balancing (RLB). Esse último utilizou uma métrica de balanceamento de carga proposta pelos autores. Em virtude dos contêineres serem migrados a partir das VMs, os resultados de consumo energia tiveram uma sobrecarga adicional de 28%. Já cada contêiner isolado representou um aumento de 1,3% no consumo de energia. Os algoritmos LC, RLB e ME apresentaram melhores resultados em relação aos demais, sendo o ME com apenas 2% a mais do que o LC e o RLB.

A pesquisa de Chen et al. (2019) utilizou o simulador de nuvem CloudSim4.0 para avaliar uma estrategia de distribuição de instâncias baseadas em contêineres dentro de VMs conforme duas listas cujos critérios de classificação foi a taxa de utilização dos recursos. O algoritmo cria uma lista ordenada que prioriza as VMs mais ociosas em detrimento das VMs menos ociosas conforme a taxa de utilização da CPU e Memória RAM (Random Access Memory). O algoritmo realiza varreduras repetitivas nas duas listas de classificação conforme o percentual uso da CPU e Memória tanto das VMs quanto dos contêineres. A estratégia aproveita a disponibilidade dos recursos nas VMs para suportar máximo possível de contêineres sem comprometer o desempenho geral da operação. De acordo com os resultados, essa estratégia otimizou a utilização dos recursos e reduziu o consumo de energia em 12,8%.

Em Yadav et al. (2018), os autores propuseram uma metodologia para comparar o desempenho das VMs VMWare e dos contêineres *Docker*. Os ambientes construídos por VMs e contêineres receberam cargas de trabalhos de um benckmark denominado sysbench que, posteriormente, forneceu resultados sobre o tempo de execução da CPU e da Memória RAM. Os experimentos tiveram variações de cinco cenários com quantidade limitadas a cinco instâncias baseadas em contêineres e VMs. O tempo de execução da memória foi avaliado através de transferências de arquivos com tamanhos variados entre 40GB, 60GB, 80GB, 100GB e 120GB. O tempo de execução da CPU foi observado através dos cálculos de números primos com intervalos variados entre 1-20.000, 1-30.000, 1-40.000, 1-50.000 e 1-60.000. Os resultados demonstraram que o desempenho dos contêineres sobre as VMs representaram diferenças médias de 1,1% e 0,69% no tempo de execução da memória e da CPU, respectivamente. Assim, a diferença do tempo de execução entre os contêineres e as VMs sobre os cenários apresentados não indicaram uma superioridade significativa de ambas tecnologias.

Bhimani et al. (2017) analisaram e compararam o desempenho de aplicativos de Biq Data executados por máquinas virtuais e contêineres. O Apache Spark consiste em um mecanismo de computação para clusters cuja finalidade é processar dados compartilhados e escaláveis em nuvens de *Biq Data*. Os *clusters*, que são formados pelos Nós do Spark, suportam a execução de várias tarefas paralelas, visando concorrer os recursos computacionais dos núcleos das CPUs, memórias e discos rígidos de todo sistema computacional. Os benchmarks denominados como K-Means, PageRank e SQLDataSource, além de outros 11 benchmarks, tiveram a função de avaliar o desempenho de aplicações classificadas como: aprendizagem de máquinas, computação de gráfica e consultas em SQL, respetivamente. As medições adotaram dois clusters distintos compostos por 8 instâncias baseadas em contêineres e VMs. Os experimentos observaram o percentual de utilização da CPU, do disco rígido e da memória em três cenários conforme a classificação do benchmark. Os experimentos de aprendizagem de máquinas aplicados aos cluster dos contêineres resultaram em 3,67%, 13% e 100% da utilização da CPU, do disco rígido e da memória, respectivamente. Ao passo que, os mesmos testes aplicados aos clusters das VMs resultaram em 16,17%, 13,83% e 71,00% da utilização da CPU, do disco rígido e da memória, respectivamente. Os experimentos de computação de gráfica aplicados aos cluster dos contêineres resultaram em 5,80%, 19,20% e 92,20% da utilização da CPU, do disco rígido e da memória, respectivamente. Enquanto, os mesmos testes aplicados aos clusters das VMs resultaram 23,80%, 23,60% e 72,80% da utilização da CPU, do disco rígido e da memória, respectivamente. Por fim, os experimentos de consultas em SQL aplicados aos cluster dos contêineres resultaram em 1,5%, 11% e 100% da utilização da CPU, do disco rígido e da memória, respectivamente. Ao passo que, os testes os mesmos aplicados aos clusters das VMs resultaram em 17,50%, 12,00% e 83% da utilização da CPU, do disco rígido e da memória, respectivamente. Os experimentos conduzidos nos contêineres apresentaram resultados melhores para a maioria dos benchmarks executados. Todavia, o benchmark K-Means demonstrou uma diferença marginal no percentual de utilização da CPU a favor das VMs. Observou-se que os testes de K-Means realizados no cluster das VMs utilizaram 2% da CPU, ao passo que o cluster dos contêineres utilizaram 5% da CPU.

Em Salah et al. (2017), os autores abordaram uma avaliação de desempenho utilizando a infraestrutura como serviço (Infrastructure as a service) da Amazon Cloud Plataform para distribuir serviços em contêineres e VMs. As duas tecnologias executaram os serviços Web dentro dos ambientes da Amazon Elastic Compute Cloud (Amazon EC2) e (Amazon ECS), respectivamente. O ambiente de experimentação foi estruturado em três cenários com variações de escopos contendo uma, duas e três instâncias de servidores Apache Web. A análise de desempenho dos serviços Web baseados em contêineres e VMs levaram em consideração as seguintes métricas: vazão, tempo de resposta e utilização da CPU. As cargas de trabalhos foram requisições enviadas simultaneamente aos servidores web durante intervalos de 20 segundos. O benchmark JMeter simulou grupos de usuários virtuais contendo 10 requisições simultâneas. Os resultados da vazão e do tempo de resposta mostraram melhores desempenhos das VMs avaliadas no três cenários. O cenário 1 observou a operação de um servidor Web carregado no contêiner e outro na VM. Os resultados da vazão e do tempo de resposta observados no cenário 1 indicaram uma diferença de 20% favor da VM. Posteriormente, o cenário 2 observou a operação dois servidores Web carregados em cada contêiner e VMs. Os resultados da vazão e do tempo de resposta observados no cenário 2 indicaram uma diferença de 50% favor da VM. Por fim, os resultados do cenário 3 apresentaram uma diferença de 27,27% a favor de VM, considerando três servidores carregados nos carregados contêineres e VMs. O percentual de utilização da CPU chegou a 100% considerando todos cenários. É importante ressaltar que tanto a Amazon EC2 quanto a Amazon ECS operam sobre máquinas virtuais gerenciadas por hypervisors. Portanto, a justificativa da inferioridade

no desempenho dos contêineres está atrelada aos *overhe*ads adicionais provocados pelas VMs que hospedaram os contêineres

O estudo de Tadesse et al. (2017) abordou uma avaliação de consumo de energia para ambientes virtuais baseados em contêineres *Docker*. O objetivo da pesquisa foi apresentar resultados sobre o percentual de utilização das CPUs e do impacto no consumo de energia quando contêineres servidores são submetidos às elevadas taxas de transferências de arquivos na rede. Os cenários experimentais variaram em até quatro instâncias baseadas em contêineres que executaram o benchmark iperf como servidores, os quais receberam cargas de trabalhos geradas por clientes que executaram o benchmark iperf. As cargas de trabalhos produziram transferências de dados que exauriram a capacidade total das interfaces de rede dos contêineres servidores. Os resultados confirmaram que o aumento de consumo de banda da rede está associado ao percentual de utilização da CPU e ao consumo de energia do computador hospedeiro. As cargas de trabalhos produziram uma utilização de quase 100% dos quatro núcleos da CPU do host servidor.

Os autores em Cuadrado-Cordero et al. (2017) fizeram uma avaliação de desempenho através da comparação entre máquinas virtuais do KVM e contêineres do Docker com ênfase na eficiência energética e na qualidade de serviço (QoS) para uma infraestrutura em nuvem. O benchmark RUBiS foi utilizado na simulação de múltiplos usuários concorrentes, de modo a paralelizar os serviços em execução nas instâncias baseadas em contêineres e máquinas virtuais de uma infraestrutura nas nuvens. O tempo de resposta foi a métrica que serviu de base para representar o QoS da rede e comparar as duas tecnologias. Os resultados mostraram que os tempos de respostas alcançados pelos serviços simultâneos em máquinas virtuais foram: 1000 ms para 15 VMs, 2000 ms para 17 VMs e 3000 ms para 19 VMs. Em contrapartida, os tempos para contêineres foram: 1000 ms para 16 contêineres, 2000 ms para 20 contêineres e 3000 ms para 24 contêineres. Para esse último caso, o resultado representou um ganho de 26%. Com relação ao consumo de energia, os resultados comprovaram que o Docker consome menos energia do que o KVM. Todavia, as medições de eficiência energética não apresentaram resultados concretos para responder as questões de custos financeiros.

Barik et al. (2016) apresentaram uma avaliação de desempenho entre virtualização e conteinerização através do Oracle Virtual Box e o Docker. O estudo comparou as duas tecnologias através de medições realizadas pelos benchmarks AIO Stress e o RAM-Speed que avaliou a largura de banda da leitura e da escrita nas memórias do tipo: cache de nível I, cache de nível II e a Random Access Memory (Memória RAM). Outro benchmark utilizado foi o IO ZONE que criou conjuntos de arquivos para gerar cargas de trabalhos no disco rígido e, consequentemente, avaliar a largura de banda na leitura e escrita. O desempenho da rede foi medido através das transferências de arquivos via os protocolos Transmission Control Protocol (TCP) e User Datagram Protocol (UDP) entre instâncias virtuais de servidores e clientes que executaram os benchmarks iperf e tbench para produzir cargas de trabalhos e mostrar relatórios da vazão da rede. O benchmark RuBBos gerou as requisições HTTP

para servidores Apache com banco de dados, e depois gerou relatórios com o tempo de conexão e a taxa máxima das requisições atendidas. A vazão em MB/s alcançada na operação de escrita da memória alocada para as duas tecnologias atingiu uma diferença de 149% a favor do contêiner. No entanto, a vazão da operação de leitura entre as duas tecnologias apresentou uma leve diferença 1,73% a favor do contêiner. O teste de desempenho da rede utilizou a interface loopback para observar o tempo gasto na transferência de um arquivo limitado a 10 gigabytes. O tempo gasto na transferência local do arquivo apresentou uma diferença de 51,74% a favor da VM. O tempo gasto para realizar o teste de requisições HTTP apresentou uma diferença de 83,04% a favor do contêiner. Assim como, o tempo gasto para realizar o teste de renderização de páginas PHP apresentou uma diferença de 112,83% a favor do contêiner. No entanto, A VM superou o contêiner em 58,01%, quando o teste de requisições HTTP incluiu uma criptografia com cifra em bloco e chaves simétricas do tipo AES (Advanced Encryption Standard) de 256 bits. No teste de emissão de certificados SSL, o contêiner emitiu um valor percentual de 126,64% a mais do que a VM. As avaliações concluíram que o contêiner superou VM na maioria dos experimentos, exceto nos testes de transferências de arquivos via interface loopback e no processamento das requisições HTTP com criptografia simétrica, nos quais das VMs apresentaram melhores resultados.

As pesquisas apresentaram resultados indicando a superioridade a favor dos contêineres quando as aplicações demandaram recursos da CPU, da memória e do disco rígido. A maioria dos experimentos também indicaram tempos de respostas menores a favor dos serviços instanciados nos contêineres. No entanto, a literatura ainda carece de trabalhos consolidados acerca da avaliação do consumo de energia e custo de operações baseadas em contêineres e VMs (Cuadrado-Cordero et al., 2017). A pesquisa conduzida por Yadav et al. (2018) executou os benchmarks visando demandar recursos da CPU e da memória alocadas para as VMs e contêineres, de modo que o fosse observado o comportamento do tempo de execução das cargas de trabalhos e a utilização dos recursos. O trabalho de Tadesse et al. (2017) apresentou um modelo de regressão linear que considerou a relação da taxa de transferência da rede com o consumo de energia e a utilização da CPU de um ambiente experimental baseado em contêineres Docker. As pesquisas Zheng et al. (2020) e Chen et al. (2019) implementaram algoritmos cujas decisões são baseadas em técnicas que otimizam as distribuições e as migrações das instâncias conforme os percentuais de ociosidade dos recursos. Os resultados indicaram que as estratégias adotadas nos algoritmos trouxeram benefícios quanto à redução do desperdício de recursos. É importante destacar que as pesquisas de Zheng et al. (2020) e Chen et al. (2019) utilizaram um simulador de nuvem denominado de Cloud-Sim4.0 para realizar os experimentos. Esta pesquisa diferencia das pesquisas Yadav et al. (2018), Tadesse et al. (2017), Zheng et al. (2020) e Chen et al. (2019) no aspecto da abrangência dos fatores e métricas avaliadas, nas variações das cargas de trabalhos e na proximidade com a operação de um ambiente real. As cargas de trabalhos e a largura de banda dos links demandaram da operação uma elevada taxa na vazão da rede, bem como altos percentuais

de utilização da CPU e consumo de energia do computador hospedeiro. A metodologia desta pesquisa realizou experimentos de simples execução e fácil replicação para que os resultados produzidos fossem de fácil entendimento. A implantação dos contêineres nas operações dos *data centers* pode minimizar o custo total de propriedade (TCO) e, consequentemente, de médio a longo prazo, oportunizar um aumento no retorno sobre o investimento (ROI¹). Sendo assim, este artigo propôs uma comparação entre a virtualização tradicional e a conteinerização através de uma avaliação desempenho com ênfase na utilização da CPU, consumo de energia e custo do kiloWatt-hora.

4 Arquitetura Experimental

Esta seção descreve o escopo da arquitetura experimental e a metodologia adotada na condução dos experimentos referentes às avaliações de desempenho e o consumo energia dos contêineres e VMs. Os experimentos utilizaram um computador hospedeiro que forneceu ambientes baseados em contêineres e VMs. Durante as medições foram coletadas amostras que observaram os comportamentos da utilização da CPU e da potência elétrica Watts demandada. O kiloWatt-hora foi a unidade de potência adotada nas exposições dos resultados de consumo de energia e custo, visto que as concessionárias de energia elétrica adotam essa unidade de energia nos cálculos de faturamento pelos serviços entregues aos consumidores. A criação dos ambientes baseados em VMs utilizou a ferramenta hypervisor KVM, ao passo que os ambientes baseados em contêineres foram criados pela ferramenta docker versão 17.03.00-ce. Sendo assim, tanto as instâncias do hypervisor KVM, quanto as instâncias do docker alocaram proporcionalmente os recursos da CPU com 4 núcleos, a memória RAM, a interface gigabit ethernet e os demais componentes do computador hospedeiro. O objetivo das medições foi assegurar a imparcialidade das comparações entre os experimentos relacionados às duas tecnologias. Os recursos alocados nas operações dos ambientes experimentais estão mais detalhados a seguir.

As medições levaram em conta três cenários com quantidades equivalentes de contêineres e VMs. Adicionalmente, os experimentos realizados nos três cenários também observaram a vazão da rede e os volumes de arquivos transferidos conforme a largura de banda suportada por cada escopo. Os experimentos realizados nos ambientes criados dentro do computador hospedeiro utilizaram benchmarks e ferramentas, que produziram cargas de trabalhos, monitoram e coletaram amostras referentes a vazão da rede, utilização da CPU e consumo de energia. Os benchmarks e ferramentas usadas neste trabalho são descritas a seguir:

 IPERF é o benchmark que foi selecionado para estressar as VMs e os contêineres através das cargas de trabalhos geradas pelos clientes de cada cenário. Adicionalmente, o iperf estabeleceu conexões entre clientes e servidores

¹ROI Return on Investment é uma relação entre o lucro líquido e o custo do investimento resultante de um investimento de alguns recursos. Em termos puramente econômicos, é uma maneira de relacionar os lucros ao capital investido.

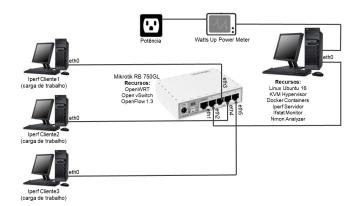


Figura 4: Topologia do ambiente de experimentação

para gerar carga na rede;

- IFSTAT consiste na ferramenta de monitoramento de tráfego de rede. A ferramenta também complementou as execuções das medições em conjunto com o iperf na coleta da vazão na rede;
- NMONANALYZER realizou o monitoramento de todos os recursos internos do computador baseado no sistema operacional Linux/Unix. Durante a geração das cargas de trabalhos pelo iperf, o nmon fez o monitoramento e a coleta de vários parâmetros dos componentes do computador hospedeiro, tais como: utilização de CPU, memória RAM, disco rígido, interface de rede, sistema de arquivos, memória virtual, etc. No entanto, este artigo focou apenas na utilização da CPU como métrica desempenho;
- POWER METER consiste na ferramenta de medição do consumo de energia cuja descrição do equipamento é a Wattsup Power Meter. Conforme mostra a Fig. 4, durante a execução dos experimentos, o computador hospedeiro teve as operações dos contêineres e VMs monitoradas pelo medidor de energia, o qual coletou amostras da potência média demandada em virtude as cargas de trabalho geradas pelos clientes iperf.

A arquitetura experimental está representada pela topologia descrita na Fig. 4. Nesta topologia, um roteador Mikrotik 750GL estabeleceu a conexão de rede com suporte a 1000Mbit/s de largura de banda entre todos os computadores. Os clientes 1, 2 e 3 estiveram conectados às portas eth3, eth4 e eth5, respectivamente. Já o computador hospedeiro esteve conectado à porta eth2. Os clientes das cargas de trabalhos possuíram as suas interfaces ethernet limitadas a uma largura de banda de 100Mbit/s. Como a interface de rede do computador hospedeiro suportava uma largura de banda de 1000Mbit/s, essa capacidade do link foi compartilhada para cada instância de contêiner ou VM cujas interfaces ethernet virtuais alocaram 100Mbit/s da interface de rede física do computador hospedeiro. Apesar desta pesquisa não focar no tema Software defined networking -SDN (RYU Project Team, 2018), o roteador implementou uma rede local baseada na arquitetura SDN sendo controlada pelo Ryu Controller (RYU Project Team, 2018).

A Tabela 1 descreve as configurações dos equipamentos usados nos experimentos. O servidor do ambiente experimental utilizou um computador Dell Inspiron 3647 com processador intel Core i5 com 4 núcleos de 3.4Ghz, memória RAM de 8GB, Disco Rígido de 1000GB e placa de rede qiqabit ethernet de 1Gbit/s. Os clientes geradores das cargas de trabalhos foram computadores da marca Hewlett-Packard (HP) com processador intel Dual Core de 2.0Ghz, memória RAM de 2GB, Disco Rígido de 500GB e placa de rede *Fast Ethernet* de 100Mbit/s. Os três computadores clientes tiverem uma instalação da distribuição Linux Ubuntu 16.10. O computador hospedeiro teve o seu disco rígido (HD) de 1000GB particionado de modo a conter duas instalações da distribuição Linux Ubuntu 16.10, sendo uma partição SDA1 de 500GB destinada à criação das VMs através do KVM e outra partição SDA2 de 500GB destinada à criação dos contêineres através do Docker.

O modo operação *TCP server* do *iperf* foi executado em cada instância de contêiner ou VM que forneceram serviços de transferências de arquivos na rede local. As demandas destinadas a esses serviços ocuparam toda largura de banda suportada pelas interfaces *ethernet* de cada cliente que executou o *iperf* no modo *TCP client*. Esse processo garantiu uma taxa elevada na vazão da rede, permitindo estressar o processamento geral do computador hospedeiro. É importante frisar que os clientes tiveram as suas interfaces *ethernet* limitadas a 100Mbit/s de largura de banda. Dessa forma, interface *gigabit ethernet* do computador hospedeiro teve disponibilidade suficiente para atender as requisições oriundas dos clientes. Os experimentos de transferências de arquivos compuseram 60 amostras com intervalos de 30 segundos entre cada observação.

A Tabela 2 descreve a alocação de todos os recursos destinados à realização dos experimentos. Estes cenários apresentaram três tipos de configurações que visaram estressar os recursos do computador hospedeiro e posteriormente resultar em amostras relativas à utilização da CPU, o consumo de energia e a vazão da rede. Quando cada contêiner ou VM era agregado para executar um experimento, o cenário expandia de capacidade de processamento conforme a quantidade de instâncias. O Escopo I e o Escopo II representaram as alternâncias dos experimentos entre as duas tecnologias dentro do próprio cenário. Os clientes das cargas de trabalhos foram utilizados à medida que uma instância de contêiner ou VM era adicionada ao experimento. Os cenários tiveram três agrupamentos de contêineres ou VMs cujas capacidades dos links da rede aumentaram as larguras de banda para 100Mbit/s, 200Mbit/s e 300Mbit/s conforme aumentava o tamanho do agrupamento para suportar um maior volume de carga de trabalho. Além disso, os escopos alternavam entre contêineres e VMs dentro de cada cenário. No Cenário I, o Escopo I utilizou apenas um cliente de carga de trabalho (Cliente 1) e uma instância baseada em contêiner (C1) para executar o iperf no modo servidor e as demais ferramentas de medição. Enquanto, o Escopo II utilizou apenas um cliente de trabalho (Cliente 1) e uma instância baseada em VM (VM1). No Cenário II, o Escopo I utilizou dois clientes (Cliente 1 e Cliente 2) para gerar as mesmas cargas de trabalhos em paralelo e duas instâncias baseadas em contêineres (C1+C2) para executar iperf no modo servidor. Seguindo raciocínio anterior, o Escopo II adotou o mesmo procedimento utilizando duas

Tabela 1: Especificações dos sistemas

Sistema Computacional	Configuração	SO	Hosts
Hardware Hospedeiro - SDA1	Intel i5, 8GB RAM, 500GB 7200RPM, Ethernet 1000 Mbit/s	Linux Ubuntu 16.10	Servidor-KVM
KVM - Máquina Virtual	Intel i5, 2GB RAM, 100GB 7200RPM, Ethernet 100 Mbit/s	Linux Ubuntu 16.10	VM1, VM2, VM3
Hardware Hospedeiro - SDA2	Intel i5, 8GB RAM, 500GB 7200RPM, Ethernet 1000 Mbit/s	Linux Ubuntu 16.10	Servidor-Docker
Docker - Contêiner	Intel i5, 2GB RAM, 100GB 7200RPM, Ethernet 100 Mbit/s	Linux Ubuntu 16.10	C1, C2, C3
Hardware - Clientes PCs	Intel Dual Core 2.0 Ghz, 2GB RAM, 500GB 5400RPM, Ethernet 100 Mbit/s	Linux Ubuntu 16.10	Cliente 1, Cliente 1, Cliente 3

Tabela 2: Números de instâncias alocadas aos experimentos

N. de Instâncias do Agregadas	Níveis de Cenários	Cargas de Trabalhos
C1	Cenário I - Escopo I	1 (Cliente 1)
VM1	Cenário I - Escopo II	1 (Cliente 1)
C1+C2	Cenário II - Escopo I	2 (Cliente 1 e Cliente 2)
VM1+VM2	Cenário II - Escopo II	2 (Cliente 1 e Cliente 2)
C1+C2+C3	Cenário III - Escopo I	3 (Cliente 1, Cliente 2 e Cliente 3)
VM1+VM2+VM3	Cenário III - Escopo II	3 (Cliente 1, Cliente 2 e Cliente 3)

instâncias baseadas em VMs (VM1+VM2). Finalmente, no Cenário III, o Escopo I utilizou três clientes que geraram as cargas de trabalhos (Cliente 1, Cliente 2 e Cliente 3) e três instâncias baseadas em contêineres (C1+C2+C3). Enquanto, o Escopo II utilizou três instâncias baseadas em VMs (VM1+VM2+VM3). Na próxima seção, os resultados dos experimentos serão apresentados com mais detalhes.

5 Resultados

Nesta seção serão apresentadas as análises e as comparações das métricas de desempenho referente aos cenários e escopos propostos nesta pesquisa. A composição das amostras se basearam nos valores médios: dos tamanhos dos arquivos transferidos em unidades de Megabytes; da vazão da rede observadas em Megabit por segundo (Mbit/s); da utilização da CPU resultante da demanda pelo serviço; e da potência consumida pela operação. Os arquivos transferidos pelo iperf aumentaram conforme a quantidade de instâncias baseadas em contêineres ou VMs que estiveram com os serviços aptos a responder as requisições dos clientes. É importante destacar que o iperf estabeleceu as conexões por meio do protocolo TCP, o qual transportou os arquivos provenientes das cargas de trabalhos. Cada escopo selecionado entre os cenários compreenderam experimentos cujos períodos de observações duraram 1800 segundos (30 minutos) divididos em 60 amostras de 30 segundos.

5.1 Cargas de Trabalhos

As intensidades das cargas de trabalhos foram caracterizadas pelos volumes de arquivos transferidos através das interfaces *ethernet* configuradas para suportar 100Mbit/s de largura de banda. Assim sendo, no Cenário I, duas medições foram realizadas, sendo uma destinada à instância baseada em contêiner e outra à instância baseada em VM, que resultaram em arquivos cujos tamanhos médios das 60 amostras corresponderam a 336,97MBytes e 337,00MBytes, respectivamente. Em seguida, o Cenário II avaliou as duas abordagens através de dois grupos distintos cujas

composições se caracterizaram por duas instâncias de contêineres e duas de VMs, sendo ambas capazes de processar as demandas relativas às duas cargas de trabalhos em paralelo, as quais resultaram em arquivos cujos tamanhos médios corresponderam a 672,37MBytes e 673,75MBytes, respectivamente. Por fim, o Cenário III adotou dois experimentos relativos às três cargas de trabalhos em paralelo que demandaram a capacidade de processamento de dois grupos compostos por três instâncias de contêineres e três de VMs que produziram amostras caracterizadas por arquivos de 1004,85MBytes e 1010,87MBytes, respectivamente.

5.2 Vazão

De acordo com a Tabela 3, considerando que os experimentos mantiveram as equivalências numéricas entre as instâncias do Escopo I e Escopo II em relação ao Cenário I, Cenário II e Cenário III. Observou-se que as médias das amostras referente aos tamanhos dos arquivos apresentaram diferenças numericamente pequenas. Por isso, devido às proximidades dos tamanhos de arquivos, pode-se afirmar que os contêineres e as VMs conseguiram processar as requisições dos clientes com desempenhos aproximados. De forma semelhante, as amostras da vazão na rede produziram taxas de transferências médias que apresentaram comportamentos semelhantes aos valores das médias relativos aos tamanhos dos arquivos resultantes das cargas de trabalhos.

Sendo assim, os volumes de MBytes transportados pelas conexões TCP estabelecidas entre os clientes e as instâncias dos servidores utilizaram as interfaces *ethernet* com capacidade para suportar 100Mbit/s de largura de banda nominal. Logo, devido à uniformidade na capacidade de transmissão das interfaces *ethernet* situadas nos contêineres e VMs, as vazões atingiram valores médios com leves diferenças entre as instâncias do Escopo I e Escopo II. Considerando que os *links* aumentaram a largura de banda para 100Mbit/s, 200Mbit/s e 300Mbit/s conforme as quantidades de instâncias agregadas nos Cenário I, Cenário II e Cenário III, respectivamente. Portanto, é importante ressaltar que as medições observaram amostras cujas vazões atingiram taxas de transferências próximos

MBytes	Cena	rio I	Cena	rio II	Cenario III		
Estatísticas	Escopo I	Escopo II	Escopo I	Escopo II	Escopo I	Escopo II	
Médias (50 amostras)	337.00	336.97	673.75	672.37	1010.87	1004.85	
Desvio Padrão	0.00	0.181	0.75	0.801	0.343	7.412	
Intervalo de Confiança (95%)	[337.00;337.00]	[336.92;337.01]	[673.56;673.94]	[672.16;672.57]	[1010.778;1010.955]	[1002.93;1006.76]	

Tabela 3: Tamanhos dos arquivos transferidos por cargas de trabalhos entre os cenários

às larguras de bandas disponibilizadas nos links estabelecidos entre os clientes e contêineres ou VMs instanciadas no computador hospedeiro. A Fig. 5 representa graficamente as comparações entre as taxas das vazões médias atingidas pelas instâncias baseadas em contêineres e VMs. As vazões atingiram altas taxas de transferência e elevados níveis de utilização da rede conforme a capacidade dos links.

Ao comparar as médias entre o Escopo I e Escopo II referente ao Cenário I, observou-se que um experimento relativo a um contêiner atingiu uma taxa de 94,10Mbit/s, enquanto que uma VM atingiu uma taxa de 96,13Mbit/s. Em seguida, o Cenário II observou que as operações paralelas relativas às duas instâncias baseadas em contêineres atingiram uma taxa média de 189,21Mbit/s. Ao passo que as operações paralelas relativas às duas instâncias baseadas em VMs atingiram uma taxa média de 192,18Mbit/s. De forma semelhante, o Cenário III observou que as operações paralelas referentes às três instâncias baseadas em contêineres atingiram uma taxa média de 284,60Mbit/s. Já as operações paralelas referentes às três instâncias baseadas em VMs atingiram uma taxa média de 286,42Mbit/s.

As vazões obtidas pelas instâncias agrupadas em todos os cenários experimentais apresentaram diferenças pequenas entre as médias. As diferenças entre as taxas médias das vazões foram de 2,03Mbit/s, 2,97Mbit/s e 1,82Mbit/s. Em termos percentuais, considerando as vazões atingidas pelos experimentos realizados nos escopos dos cenários I, II e III, as VMs superaram os contêineres em 2,16%, 1,57% e 0,64%, respectivamente. Seguindo o mesmo raciocínio, todavia, a favor do contêineres, as diferenças percentuais entre as médias dos arquivos transferidos foram 0,01%, 0,21% e 0,6%, respectivamente. A fim de verificar se as diferencas são estatisticamente significante, testes estatísticos foram realizados. Primeiramente, foi verificado a normalidade das amostras, de modo que o teste de Shapiro-Wilk (SHAPIRO and WILK, 1965) avaliou a conformidade das médias com uma distribuição normal com um nível de significância de 0,05 (5%). Os valores das probabilidades de significâncias (valor-p) foram inferiores ao nível de significância de 0,05. Sendo assim, as médias não seguem uma distribuição normal. Uma vez que as amostras não seguem uma distribuição normal, o teste não-paramétrico chamado de Wilcoxon-Mann-Whitney (Mann and Whitney, 1947) foi adotado para comparar os escopos com um nível de significância de 0,05 (5%). Os resultados dos testes aplicados em todas as amostras indicaram uma probabilidade de significância (valor-p) inferior ao nível de significância de 0,05. Dessa forma, o resultado do teste revela que existem diferenças estatisticamente significante entre os escopos comparados. Todavia, no ponto de vista dos desempenhos dos contêineres e VMs, as evidências empíricas apontaram que as diferenças entre as médias das taxas apresentaram valores mínimos, visto que, a favor dos contêineres, as diferenças entres as médias dos arquivos transferidos foram 0,03MBytes, 1,38MBytes e 6,02MBytes, respectivamente.

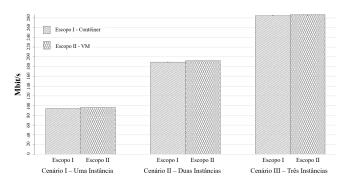


Figura 5: As taxas de transferências relativas às vazões produzidas por links de 100 Mbit/s, 200 Mbit/s e 300 Mbit/s de largura de banda

Os resultados dos experimentos demonstraram que as taxas médias das vazões e os tamanhos médios dos arquivos transferidos na rede apresentaram crescimentos proporcionais aos números de clientes nas cargas de trabalhos e os quantitativos de instâncias baseadas em contêineres e VMs utilizadas nos Cenário I, cenario II e Cenário III. Nesse caso, as taxas médias das vazões dos links aumentaram nas proporções de 94,10Mbit/s e 96,13Mbit/s em virtude da inclusão de uma instância baseada em contêiner ou uma baseada em VM, respectivamente. Além disso, os volumes de arquivos transferidos com base na vazão aumentaram na proporção média de 337,00 MBytes para cada cliente, contêiner ou VM inseridos nos experimentos.

Utilização da CPU

Basicamente, as proporções dos volumes de MBytes e largura de banda dos links inerente aos experimentos produziram efeitos práticos no desempenho geral do computador hospedeiro. As medições realizadas através da ferramenta NMON Analyzer coletou amostras oriundas das cargas de trabalhos tratadas pela CPU que foi compartilhada entre todas as instâncias baseadas em contêineres e VMs. As análises das amostras levaram em consideração os valores médios das variáveis de ambientes responsáveis pelo monitoramento do desempenho do hardware da CPU. Os níveis de utilização e ociosidade da CPU foram observados através de quatro variáveis de ambiente representadas por valores percentuais. Os resultados representados na Fig. 6 (a) mostraram que os contêineres dos escopos I apresenta-

Utilização	User%			Sys%			Idle%			Busy%			
QTD de instâncias	Média	SD	IC	Média	Média SD IC		Média	édia SD IC		Média	SD	IC	
Cenário I - Escopo I	1.23	0.19	[1.19;1.28]	4.48	0.25	[4.41;4.54]	94.16	0.32	[94.07;94.24]	5.71	0.32	[5.63;5.80]	
Cenário I - Escopo II	6.09	0.42	[5.99;6.20]	23.55	0.40	[23.45;23.65]	70.02	0.49	[69.90;70.14]	29.64	0.49	[29.51;29.77]	
Cenário II - Escopo I	1.56	0.57	[1.41;1.70]	7.09	0.43	[7.0;7.2]	91.21	0.72	[91.03;91.40]	8.64	0.74	[8.45;8.83]	
Cenário II - Escopo II	13.22	0.44	[13.12;13.34]	47.53	2.40	[46.91;48.15]	38.98	2.50	[38.34;39.62]	60.75	2.50	[60.11;61.40]	
Cenário III - Escopo I	2.27	1.02	[2.00;2.53]	8.5	0.46	[8.38;8.62]	89.10	1.06	[88.83;89.34]	10.77	1.06	[10.50;11.04]	
Cenário III - Escopo II	22.23	1.49	[21.84;22.61]	65.07	1.56	[64.66;65.47]	12.58	0.78	[12.38;12.78]	87.3	0.77	[87.1;87.5]	

Tabela 4: Percentual de ociosidade (idleness) e utilização da CPU entre os cenários

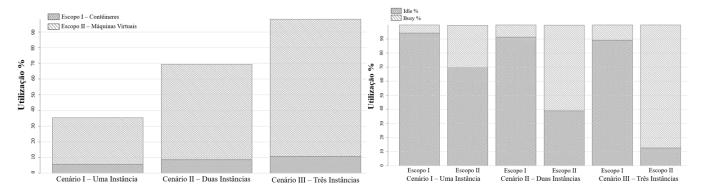


Figura 6: Comparação da utilização da CPU entre os cenários baseados em Contêineres e VMs: (a) Percentual de utilização da CPU dos escopos I e II - (b) Percentual de utilização das variáveis Idle e Busy

ram percentuais de utilização da CPU consideravelmente inferiores quando comparados às VMs dos escopos II.

A Tabela 4 apresenta os resultados estatísticos que consideraram as médias aritméticas, os desvios padrão (SD) e os intervalos de confiança (IC) de 95% das variáveis de ambiente a seguir: *User*, *Sys*, *Idle* (ociosidade) e *Busy* (ocupação). A varável *User* é definida pelo percentual de utilização da CPU em conformidade com as sessões abertas pelos usuários. Nesse caso, apenas um usuário esteve logado no SO durante as medições. A variável *Sys* representa o percentual de utilização da CPU por aplicações e serviços do SO ou *Kernel*. A variável *Idle* representa o percentual de ociosidade da CPU. Por fim, a varíavel *Busy* significa o percentual total de ocupação da CPU.

As medições demonstraram que tanto os contêineres quanto as VMs sem processar as cargas de trabalhos não representaram impacto na utilização CPU, independentemente da quantidade de instâncias utilizadas nos três cenários experimentais. Em virtude da ausência de cargas de trabalhos destinadas aos servidores instanciados pelos contêineres e VMs, a ociosidade da CPU permaneceu próxima de 99,8%. É importante destacar que o comportamento da variável *Idle* é inversamente proporcional ao comportamento da variável *Busy* cujo resultado pode ser obtido pelo somatório das variáveis *User* e *Sys*. A análise de desempenho deste trabalho focou nas variáveis *Busy* e *Idle*, visto que ambas representam quase a totalidade do processamento no ambiente *Linux/Unix*.

A Fig. 6 (b) mostra que a variável sys representou uma maior carga na utilização da CPU em virtude do *Hypervisor* ter executado as chamadas de sistema direcionadas aos processos que compõem o *kernel* do sistema hospedeiro. O *Hypervisor* intermedeia o processo de conversão das instruções entre o sistema convidado e o sistema hos-

pedeiro. Por outro lado, a variável *user* representou uma menor carga na utilização CPU, porque os seus processos são executados no espaço do usuário. Dessa forma, em razão das cargas de trabalhos e dos agrupamentos de VMs do escopo II, o *Hypervisor KVM* e os sistemas operacionais convidados representaram uma sobrecarga adicional no percentual da variável sys.

As cargas de trabalhos geradas pelo iperf observou-se que uma instância baseada em contêiner (C1) e uma baseada em VM (VM1) representaram 5,71% e 29,64% de utilização média da CPU, respectivamente. Ao passo que a ociosidade (Idle) da CPU correspondeu a 94,16% para uma instância baseada em contêiner e 70,02% para uma baseada em VM. Em seguida, o Cenário II compreendeu dois escopos experimentais caracterizados pelos paralelismos de duas instâncias baseadas em contêineres e duas baseadas em VMs. Os experimentos submetidos aos dois contêineres (C1+C2) e as VMs (VM1+VM2) resultaram no percentual de utilização de 8,64% e 60,75%, respectivamente. Consequentemente, a ociosidade (Idle) da CPU relativa às duas instâncias simultâneas baseadas em contêineres e VMs resultaram nas médias de utilização de 91,21% e 38,98%, respectivamente. Por fim, a utilização da CPU relativa aos três contêineres (C1+C2+C3) e as VMs (VM1+VM2+VM3) resultaram nas médias 10,77% e 87,3%, respectivamente. Já a ociosidade da CPU relativa às três instâncias simultâneas baseadas em contêineres e VMs resultaram nas médias de 89,10% e 12,58%, respectivamente.

Os resultados provenientes das análises demonstraram uma sobrecarga drasticamente inferior na utilização da CPU pelos contêineres em comparação às VMs, levando em conta todos os cenários. É importante destacar que os contêineres *Docker* são empacotados em imagens contendo

apenas arquivos, dependências e bibliotecas necessárias ao funcionamento da aplicação, bem como são carregados sobre o kernel do sistema hospedeiro. Sendo assim, os escopos baseados em contêineres apresentaram leves cargas na utilização da CPU, porque a tecnologia dispensa a sobrecarga adicional do Hypervisor e do sistema operacional convidado. Isto é, uma VM (VM1) do Cenário I representou um acréscimo de 23,42% de sobrecarga na utilização da CPU em relação a um contêiner (C1). Em seguida, considerando os resultados do Cenário II, o impacto das duas VMs (VM1+VM2) na utilização da CPU representaram um acréscimo de 52,21% de sobrecarga em comparação aos dois contêineres (C1+C2). Por fim, as três VMs (VM1+VM2+VM3) do Cenário III representaram um acréscimo de 76,53% de sobrecarga na utilização da CPU em comparação aos contêineres (C1+C2+C3). Portanto, o Hypervisor, as VMs e os sistemas operacionais convidados utilizados nos escopos II dos cenários I, II e III geram em torno de 23,42%, 52,21% e 76,53% de sobrecargas adicionais na utilização da CPU, respectivamente. Além disso, os contêineres apresentaram um melhor equilíbrio na utilização na CPU à medida que um contêiner era adicionado, a sobrecarga da CPU aumentava levemente em aproximadamente 1,82%, ao passo que, cada VM adicionada aumentava a sobrecarga da CPU em 28,89%.

5.4 Consumo de Energia

Os primeiros experimentos levaram em consideração um período de ausência de cargas de trabalhos destinadas aos contêineres e VMs. Dessa forma, o estado de ociosidade do computador hospedeiro foi observado pelo medidor watts up power meter que forneceu amostras da demanda de potência média. Durante um intervalo de tempo, a medição da energia consumida observou a ausência de carga de trabalho direcionada à operação fornecida pelo computador hospedeiro. As coletas forneceram 60 amostras contendo a potência média demandada por cada cenário e escopo. As potências médias serviram de referência para derivar novas medidas de consumo de energia relacionando algumas variações de tempo e uma tarifa média cobrada por kWh. O tempo observação do estado de ociosidade teve uma duração de 30 minutos e cada coleta de amostra considerou intervalos de 30 segundos. A relação entre a potência média e a utilização da CPU foram representadas graficamente com base nos resultados oriundos das cargas de trabalhos.

As amostras inerentes aos consumos de energia apresentaram uma potência média de 27,02 *Watts* que estabeleceu uma linha de base, chamada de valor de referencia (VR), para referenciar as potências adicionais em decorrência dos agrupamentos de contêineres e VMs adotados nos cenários experimentais. A energia consumida foi calculada de acordo com a Eq. (1), onde a variável P representou a potência (W) demandada pelo computador hospedeiro no estado de ociosidade. A constante 1000 foi empregada na conversão da unidade de potência Watt para a unidade kilowatt (kW). A variável $\triangle t$ representou as variações de tempos com períodos relacionado às horas, dias, meses e anos. Neste estudo, as análises de consumo de energia e custo estabeleceram os seguintes períodos: 24hs, 30 dias e 12 meses.

A aplicação da Eq. (1) se baseou na média 27,02W em substituição à variável P para derivar o valor médio em kilowatt (kW). O resultado da conversão pela fração correspondeu a 0,02702kW. Essa nova unidade serviu de multiplicador para as variações de tempos substituídos na variável $\triangle t$. Os cálculos da Eq. (1) adotaram unidades de tempos com períodos variando em horas. Os períodos adotados nas análises de consumo de energia corresponderam a 24h; 720h (24 x 30); e 8,640h (24 x 30 x 12). Sendo esses valores equivalentes a um dia, um mês e um ano, respectivamente. Logo, os resultados relativos ao consumo de energia corresponderam a 0,648kW/dia; 19,45kW/mês; e 233,45kW/ano. Este estudo tratou como potência adicional (PA) qualquer consumo de energia gerado a partir do valor de referência (VR ou ocioso) e a potência total (PT) foi derivada da soma entre PA e VR. É importante salientar que o valor da PA tende a crescer em decorrência do acréscimo de instâncias aptas a processar as cargas de trabalhos.

$$\triangle E(kWh) = \frac{P}{1000}.\triangle t \tag{1}$$

A Tabela 5 representa o consumo de energia em razão das instâncias agregadas ao estado de ociosidade do computador hospedeiro. Então, em virtude das amostras conferidas no Cenário I, a PA consumida pelo contêiner C1 acrescida da potência consumida pelo estado de ociosidade resultou na PT dada pela soma a seguir: 2,38W (PA) + 27,02W (ocioso) = 29,40W (PT). Já a PA consumida pela VM1 acrescida da potência consumida pelo estado de ociosidade resultou na PT dada pela soma a seguir: 15,77W (PA) + 27,02W (ocioso) = 42,79W (PT). Consequentemente, as amostras conferidas às operações dos contêineres C1+C2 do Cenário II resultaram nas PA e PT dadas pelos valores: 3,00W (PA) e 30,02W (PT). No caso das operações das VM1+VM2 do Cenário II resultaram nas PA e PT dadas pelos valores: 22,13W (PA) e 49,15W (PT). Por fim, as amostras conferidas aos contêineres C1+C2+C3 e às VM1+VM2+VM3 do Cenário III resultaram nas PA e PT dadas pelos valores 4,68W (PA) e 31,70W (PT); 27,96W (PA) e 54,98W (PT), respectivamente.

Sob outra perspectiva, as análises posteriores consideraram a variação do tempo e o consumo de energia em relação ao mês e ano. Sendo assim, a análise do Cenário I incluiu apenas uma instância baseada no contêiner C1 e uma baseada na VM VM1 da operação com intuito de calcular a energia consumida de ambas as instâncias isoladamente. Considerando a operação do contêiner C1 no Cenário I, os resultados da PA e PT foram dados pelos valores 1,71kW/mês e 21,17kW/mês; 20,53kW/ano e 253,99kW/ano. Da mesma forma, a operação da VM1 produziu os resultados da PA e PT dados pelos valores 11,36kW/mês e 30,81kW/mês; 136,27kW/ano e 369,72kW/ano. Em termos percentuais do consumo de energia do Cenário I, a operação do contêiner C1 e da VM1 demandaram uma potência adicional de aproximadamente 8,80% e 58,4%, respectivamente.

Considerando a análise de consumo de energia do Cenário II, as operações das duas instâncias paralelas dos contêineres C1+C2 produziram os resultados da PA e PT dados pelos valores: 2,16kW/mês e 21,61kW/mês; 25,92kW/ano e 259,37kW/ano. Ao passo que as operações das duas instân-

_	•			,		•					
OTD de instâncias	Potência	Potência Total (W)	Estatísticas		Potência	Potência	Potência	Potência	Potência	Potência	
Q1D uc instancias	Adicional (W)		IC Watts	SD Watts	Adicional (kW/dia)	Total (kW/dia)	Adicional (kW/mês)	Total (kW/mês)	Adicional (kW/ano)	Total (kW/ano)	
Cenário I - Escopo I	2.38	29.40	[29.36;29.43]	0.12	0.057	0.71	1.71	21.17	20.53	253.99	8.8%
Cenário I - Escopo II	15.77	42.79	[42.73;42.85]	0.24	0.379	1.03	11.36	30.81	136.27	369.72	58.4%
Cenário II - Escopo I	3.00	30.02	[29.96;30.07]	0.21	0.072	0.72	2.16	21.61	25.92	259.37	11.1%
Cenário II - Escopo II	22.13	49.15	[49.05;49.23]	0.35	0.531	1.18	15.93	35.38	191.16	424.61	81.9%
Cenário III - Escopo I	4.68	31.70	[31.61;31.79]	0.35	0.112	0.76	3.37	22.83	40.45	273.90	17.3%
Cenário III - Escopo II	27.96	54.98	[54.93;55.03]	0.18	0.671	1.32	20.13	39.59	241.59	475.04	103.5%

Tabela 5: Relação do consumo de energia considerando a quantidade de instâncias baseadas em contêineres e VMs

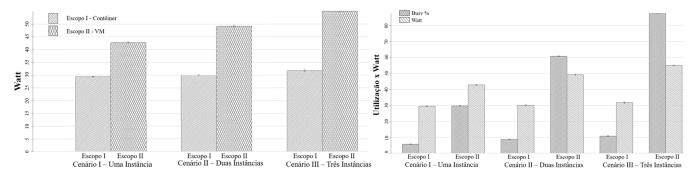


Figura 7: Relação Utilização da CPU e Consumo de energia dos cenários baseados em VMs e contêineres: (a) Potência demandada pelos escopos - (b) Relação Utilização (Busy) x Potência (Watt)

cias paralelas das VM1+VM2 produziram os resultados da PA e PT dados pelos valores 15,93kW/mês e 35,38kW/mês; 191,16kW/ano e 424,61kW/ano. Em termos percentuais do consumo de energia do Cenário II, as operações dos contêineres C1+C2 e das VM1+VM2 demandaram uma potência adicional de aproximadamente 11,1% e 81,9%, respectivamente.

Na análise de consumo de energia do Cenário III, as operações das três instâncias paralelas dos contêineres C1+C2+C3 produziram os resultados da PA e PT dados pelos valores 3,37kW/mês e 22,83kW/mês; 40,45kW/ano e 273,90kW/ano. Já as operações das três instâncias paralelas das VM1+VM1+VM3 produziram os resultados da PA e PT dados pelos valores 20,13kW/mês e 39,59kW/mês; 241,59kW/ano e 475,04kW/ano. Em termos percentuais do consumo de energia do Cenário III, as operações dos contêineres C1+C2+C3 e das VM1+VM2+VM3 demandaram uma potência adicional de aproximadamente 17,3% e 103,5%, respectivamente. A Fig. 7 (a) representa graficamente os valores médios das energias consumidas em decorrência das demandas destinadas às instâncias agrupadas nos cenários experimentais.

A segunda análise de consumo de energia conferiu as diferenças percentuais em virtude do acréscimo de instâncias na operação. Para este propósito, a subtração efetivada entre o valor da potência consumida pela operação atual e a operação posterior mostraram os valores acrescidos em decorrência do aumento do nível da operação do Cenário II para o Cenário II e, consequentemente, do Cenário II para o Cenário III. Dessa forma, observou-se que a diferença entre a capacidade da operação oferecida pelo contêiner C1 subtraída da capacidade da operação dos contêineres C1+C2 correspondeu a um aumento de 0,62W ou 2,12%. Enquanto a diferença entre a capacidade da operação oferecida pelos contêiner C1+C2 subtraída da capacidade da operação dos contêineres C1+C2+C3 correspondeu a um aumento de 1,68W ou 5,6%. Considerando os cenários dos

escopos baseados em VMs, observou-se que a diferença entre a capacidade da operação oferecida pela VM1 subtraída da capacidade da operação das VM1+VM2 correspondeu a um aumento de 6,35W ou 14,85%, enquanto a diferença entre a capacidade da operação oferecida pelas VM1+VM1 subtraída da capacidade da operação das VM1+VM2+VM3 correspondeu a um aumento de 5,84W ou 11,88%. Esses resultados indicaram que as operações baseadas em VMs apresentaram consumo de energia mais acentuados do que contêineres.

De acordo com a Fig. 7 (b), o relacionamento entre os fatores potência (Watt) e a utilização da CPU (Busy) cresceram conforme a capacidade de operação dos cenários I, II e II suportaram maiores volumes de cargas de trabalhos. No entanto, os cenários baseados em contêineres entregaram os serviços com menores níveis de utilização da CPU e potência consumida. Considerando os Escopos I de todos os cenários, as cargas de trabalhos demandaram praticamente um terço dos recursos do computador hospedeiro e o consumo de energia apresentou uma variação de potência bem equilibrada. Na perspectiva do melhor aproveitamento dos recursos computacionais, os resultados apresentados no Escopo II do Cenário III descartaria a possibilidade da inclusão de uma nova instância baseada em VM, visto que os níveis de utilização da CPU estão próximos da saturação. Nesse caso, seria necessário ampliar a capacidade computacional da infraestrutura ou fazer novas aquisições de recursos para balancear as cargas de trabalhos. Em contrapartida, os resultados apresentados no Escopo I do Cenário III permitiria a inclusão de novas instâncias baseadas em contêineres, visto que o computador hospedeiro possui disponível cerca de dois terços da sua capacidade de processamento. Além do mais, a economia de energia do Escopo I em relação ao Escopo II do Cenário III foi de 23,28W ou 73,43% a favor dos contêineres.

5.5 Custo do Consumo de Energia

A análise de consumo de energia possibilita um maior controle e gestão nos gastos com energia elétrica dos data centers. Além do mais, uma boa prática de economia de energia promove uma cultura de preservação do meio ambiente e ainda por cima proporciona um melhor custo-benefício de médio e a longo prazo na utilização de recursos computacionais. Assim, os resultados obtidos dos consumos de energia permitem realizar uma análise de custo com base no valor médio da tarifa cobrada por kWh. De acordo com agência nacional de energia elétrica (Aneel, 2019), a tarifa média nacional de energia por kilowatt-hora custa em torno de R\$ 0,57. Sendo assim, as projeções de custos deste trabalho levaram em conta uma tarifa de R\$ 0,57 por kWh como referência para obter os resultados apresentados na Tabela 5. Para tanto, a Eq. (2) foi uma extensão da Eq. (1) com a inclusão da varável Tarifa que representa o preço do kWh cobrado pela concessionária de energia elétrica (dado em R\$).

Conforme apresentado anteriormente, o consumo de energia da operação do computador hospedeiro no estado ocioso correspondeu às seguintes médias: 0,648kW/dia, 19,45kW/mês e 233,45kW/ano. Com a aplicação da Eq. (2) nas três médias anteriores, os valores de custos médio referente ao consumo de energia foram os seguintes: R\$ 0,372/dia, R\$ 11,16/mês e R\$ 133,96/ano, respectivamente. Então, caso o computador hospedeiro permanecesse operando no estado de ociosidade durante 30 dias ou 12 meses, os valores cobrados com base na tarifa seriam, respectivamente: R\$ 11,16/mês e R\$ 133,96/ano. O custo adicional se refere ao valor do custo atribuído a uma instância e o custo total é o resultado da soma entre o custo adicional e o custo associado ao estado ocioso. Portanto, a operação do Cenário I atribuída ao contêiner C1 produziu um custo adicional de R\$ 0,98/mês e R\$ 11,78/ano ou um aumento adicional de 8,8% no custo do estado ocioso. Em contrapartida, a operação atribuída à VM1 produziu um custo adicional mais acentuado de R\$ 6,52/mês e R\$ 78,19/ano, representando um aumento adicional de 58,4% no custo do estado ocioso.

$$\triangle Custo(kWh) = \frac{P}{1000}.\triangle t.Tarifa$$
 (2)

As operações fornecidas pelas duas instâncias baseadas em contêineres C1+C2 produziram um custo adicional de R\$ 1,24/mês e R\$ 14,87/ano ou um aumento de 11,1% no custo do estado ocioso. Por outro lado, as operações das duas instâncias baseadas em VM1+VM2 produziram um custo adicional de R\$ 9,14/mês e R\$ 109,69/ano, representando um aumento de 81% em relação ao estado ocioso. Por fim, as operações fornecidas pelas três instâncias baseadas em contêineres C1+C2+C3 produziram um custo adicional de R\$ 1,93/mês e R\$ 23,21/ano ou um aumento de 17,3% em comparação ao custo do estado ocioso. Em compensação, as VM1+VM2+VM3 produziram um custo adicional de R\$ 11,55/mês e R\$ 138,63/ano, representando um aumento de 103,5%. Esse último caso ultrapassou duas vezes a mais o custo relativo ao estado ocioso.

Em suma, as operações oferecidas pelos contêineres dos

Escopos I de todos os cenários (C1; C1+C2; C1+C2+C3) representaram os custos mensais na ordem de R\$ 12,15/mês, R\$ 12,40/mês e 13,10/mês, respectivamente. De forma semelhante, os custos anuais para manter as operações dos contêineres de todos os cenários (C1; C1+C2; C1+C2+C3) foram de R\$ 145,75/ano, R\$ 148,84/ano e 157,17/ano. Por outro lado, as operações oferecidas pelas VMs dos Escopos II de todos os cenários (VM1; VM1+VM2; VM1+VM2+VM3) representaram custos mensais na ordem de R\$ 17,68/mês, R\$ 20,30/mês e 22,72/mês, respectivamente. Da mesma forma, os custos anuais para manter as operações das VMs de todos cenários (VM1; VM1+VM2; VM1+VM2+VM3) representaram os valores na ordem de R\$ 212,16/ano, R\$ 243,66/ano e 272,59/ano, respectivamente. Os resultados indicam que os percentuais de custos para cada VM agregada à operação são mais acentuados em relação a cada contêiner agregado à operação.

Uma outra análise avaliou o quão mais onerosa uma operação seria em relação à outra. Para tanto, as diferenças dos custos associados às operações entre os escopos de cada cenário foram calculados. Dessa forma, as diferenças dos custos das operações dos contêineres e VMs foram calculadas utilizando a substração dos custos vinculados às instâncias das mesmas tecnologia adotados nos dois escopos dos três cenários, ou seja, a diferença entre o custo anual (R\$ 14,87) do Escopo I no Cenário II (C1+C2) subtraído pelo custo anual (R\$ 11,78) Escopo I no Cenário I (C1). A diferença do custo anual (R\$ 23,21) do Escopo I do Cenário III (C1+C2+C3) subtraído pelo custo anual (R\$ 14,87) Escopo I do Cenário II (C1+C2). As diferenças entre as operações dos cenários baseados em contêineres resultaram nos valores R\$ 3,09 e R\$ 8,34, respectivamente. Da mesma forma, a diferença do custo anual (R\$ 109,69) gerada pela operação das VMs do escopo II no Cenário II (VM1+VM2) subtraído pelo custo anual (R\$ 78,19) do Escopo II no Cenário I (VM1). A diferença do custo anual (R\$ 138,63) gerada pela operação Escopo II no Cenário III (VM1+VM2+VM3) subtraído pelo custo anual (R\$ 109,69) do Escopo II no Cenário II (VM1+VM2). As diferenças entre as operações dos cenários baseados em VMs resultaram nos valores R\$ 31,50 e R\$ 28,94, respectivamente. Logo, os resultados referentes aos custos atribuídos às operações dos contêineres representaram um impacto financeiro bem inferior em relação às operações das VMs.

Outro ponto observado se referiu à redução do custo para manter as operações em funcionamento. Considerando que ocorressem migrações das operações oferecidas pelas VMs para as operações baseadas contêineres. As reduções dos custos foram calculados através das diferenças dos custos gerados pelos contêineres do Escopo I menos os custos das VMs do Escopo II. Considerando as operações do Cenário I, o custo anual (R\$ 78,19) atribuído à VM1 subtraído pelo custo anual (R\$ 11,78) do contêiner C1. De forma semelhante, as operações do Cenário II, o custo anual (R\$ 106,69) atribuído às VM1+VM2 subtraído pelo custo anual (R\$ 14,87) dos contêineres C1+C2. Por fim, o custo anual (R\$ 138,63) atribuído às VM1+VM2+VM3 subtraído pelo custo anual (R\$ 23,21) dos contêineres C1+C2+C3. Considerando os resultados dos três cenários anteriores, as economias geradas em decorrência das migrações das operações baseadas em VMs para contêineres assegurariam reduções de custos correspon-

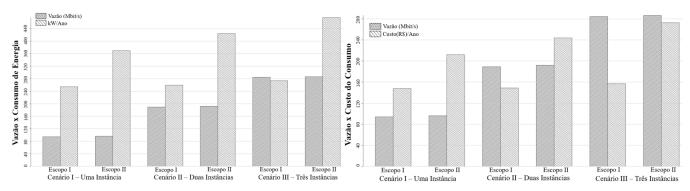


Figura 8: Relação da demanda (Vazão) pelo serviço com o consumo de energia dos escopos baseados em contêineres e VMs: (a) Vazão x kW/Ano - (b) Vazão x Custo/Ano

dente a R\$ 66,41/ano, R\$ 94,82/ano e R\$ 115,42/ano, respectivamente. Do ponto de vista da redução do custo, uma elevada quantidade de VMs migradas para contêineres poderia representar uma redução considerável no custo do consumo de energia atribuído à operação do data center.

Um aspecto relevante pontuado na Fig. 8 (a) é a relação entre a vazão da rede (Mbit/s) e o consumo de energia. Apesar das instâncias dos Escopos I e II terem atingido taxas de transferências com pouca disparidade, observouse que os tempos computacionais das VMs dos Escopos II demandaram mais energia do que os contêineres dos Escopos I. Sendo assim, independentemente do valor da tarifa cobrada pela concessionária de energia elétrica, a discrepância no consumo de energia das VMs refletirá diretamente no custo operacional da infraestrutura. A Fig. 8 (b) mostra a relação entre a vazão da rede (Mbit/s) e o custo no consumo de energia derivado a partir da tarifa de referência aplicada na Eq. (2). Os contêineres dos Escopos I representaram uma redução de custo na ordem de: 43,59% para o Cenário I; 63,71% para o cenário II; e 73,44% para o Cenário III. Portanto, a adoção de ambientes baseados em contêineres torna-se uma alternativa viável para minimizar o custo operacional da infraestrutura e contribui diretamente na redução da emissão de gases de efeitos estufa.

6 Conclusões e Trabalhos Futuros

Este artigo abordou um estudo sobre avaliação de desempenho das operações provisionadas por abordagens denominadas de conteinerização e virtualização. Para esse propósito, uma infraestrutura computacional compartilhou parcelas equiparadas de hardware com intuito de observar o comportamento da vazão na rede, da utilização da CPU e do custo do consumo de energia em virtude de operações baseadas nos contêineres distribuídos pelo docker engine e máquinas virtuais distribuídas pelo hypervisor KVM. Os serviços distribuídos pelos contêineres e VMs foram avaliados por cargas de trabalhos caracterizadas por volumes de MBytes transferidos conforme as capacidades das vazões atingidas pelas instâncias distribuídas nos cenários experimentais.

Os experimentos observados nas operações de transferências de arquivos providas pelos contêineres e VMs produziram amostras cujos volumes de MBytes e taxas de transferências apresentaram resultados com leves disparidades entre as médias dos Escopos I e Escopos II. Nos Cenários I, II e II, as taxas das vazões atingidas pelas VMs superaram os contêineres em 2,16%, 1,57% e 0,64%, respectivamente. Todavia, a favor do contêineres, as diferenças percentuais entre as médias dos arquivos transferidos foram 0,01%, 0,21% e 0,6%, respectivamente. Esses resultados indicaram que uma provável migração de um ambiente baseado em VMs para um baseado em contêineres preservaria a infraestrutura existente, visto que as operações de transferência de arquivos e a vazão da rede providas pelos contêineres não sofreriam degradações de desempenhos em decorrência das cargas de trabalhos suportadas pelas VMs.

Além disso, considerando a análise de desempenho da utilização da CPU e consumo de energia entre as operações providas pelos contêineres e VMs, os resultados revelaram uma vantagem considerável dos contêineres para os cenários avaliados. A utilização da CPU observada nas operações dos contêineres apresentaram valores percentuais bem inferiores em relação às operações das VMs. Da mesma forma, a energia consumida nas operações dos contêineres geraram reduções nos custos por kWh nas ordens de 45,56%, 63,71% e 73,44%. Neste caso, os resultados confirmaram uma maior eficiência energética dos contêineres.

Os trabalhos futuros desta pesquisa irão avaliar um estudo de caso utilizando uma aplicação WEB com instâncias replicadas em *clusters* compostos por contêineres e VMs. Os experimentos serão realizados por cargas de trabalhos arbitradas por níveis de requisições previamente definidos. As medições observarão a utilização da CPU, o consumo de energia e o tempo de resposta do ambiente experimental. Em seguida, a pesquisa assumirá uma fase de modelagem com ênfase nas Redes de Petri Estocástica que vão poder auxiliar no planejamento da capacidade de cenários com níveis de instâncias mais elevados e cargas de trabalhos controladas. Em outras palavras, as Redes de Petri Estocástica ajudarão a predizer os comportamentos das métricas através de representações de cenários.

Agradecimentos

Esta pesquisa foi parcialmente financiada pelo CNPq - Brasil, processo 406263/2018-3.

Referências

- Andrae, A. and Edler, T. (2015). On global electricity usage of communication technology: Trends to 2030, *Challenges* 6: 117–157. http://dx.doi.org/10.3390/challe6010117.
- Aneel (2019). Agência Nacional de Energia Elétrica. Disponpivel em http://www.aneel.gov.br/ranking-das-tarifas.
- Barik, R. K., Lenka, R. K., Rao, K. R. and Ghose, D. (2016). Performance analysis of virtual machines and containers in cloud computing, 2016 International Conference on Computing, Communication and Automation (ICCCA), pp. 1204–1210. http://dx.doi.org/10.1109/CCAA.2016.7813925.
- Belkhir, L. and Elmeligi, A. (2018). Assessing ict global emissions footprint: Trends to 2040 & recommendations, *Journal of Cleaner Production* **177**: 448 463. https://doi.org/10.1016/j.jclepro.2017.12.239.
- Bhimani, J., Yang, Z., Leeser, M. and Mi, N. (2017). Accelerating big data applications using lightweight virtualization framework on enterprise cloud, 2017 IEEE High Performance Extreme Computing Conference (HPEC), pp. 1–7. http://dx.doi.org/10.1109/HPEC.2017.8091086.
- Chen, F., Zhou, X. and Shi, C. (2019). The container deployment strategy based on stable matching, 2019 IEEE 4th International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), pp. 215–221. http://dx.doi.org/10.1109/ICCCBDA.2019.8725707.
- Chirammal, H., Mukhedkar, P. and Vettathu, A. (2016). Mastering KVM Virtualization, Packt Publishing.
- Cisco (2019). BWorld Robot Control Software. Disponível em https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.html#_Toc532256803.
- Cuadrado-Cordero, I., Orgerie, A. and Menaud, J. (2017). Comparative experimental analysis of the quality-of-service and energy-efficiency of vms and containers' consolidation for cloud applications, 2017 25th International Conference on Software, Telecommunications and Computer Networks (SoftCOM), pp. 1–6. http://dx.doi.org/10.23919/SOFTCOM.2017.8115516.
- Deka, G. C. (2014). Cost-benefit analysis of datacenter consolidation using virtualization, *ITProfessional* **16**(6): 54–62. http://dx.doi.org/10.1109/MITP.2014.89.
- Delforge, P. and Whitney, J. (2014). Data center efficiency assessment, Natural Resources Defense Council (NRDC) pp. 1-35. Disponível em https://www.nrdc.org/sites/default/files/data-center-efficiency-assessment-IP.pdf.
- Ivanov, K. (2017). KVM Virtualization Cookbook, Packt Publishing.
- Jones, N. (2018). How to stop data centres from gobbling up the world's electricity, *Springer Nature Limited* **561**: 163–166. http://dx.doi.org/10.1038/d41586-018-06610-y.

- Li, Z., Kihl, M., Chen, Y. and Zhang, H. (2018). Two-Stage Performance Engineering of Container-based Virtualization, Advances in Science, Technology and Engineering Systems Journal 3(1): 521–536. http://dx.doi.org/10. 25046/aj030163.
- Mann, H. B. and Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other, *Ann. Math. Statist.* **18**(1): 50–60. https://doi.org/10.1214/aoms/1177730491.
- Nickoloff, J. and Kuenzli, S. (2019). *Docker in Action*, Manning Publications.
- Raho, M., Spyridakis, A., Paolino, M. and Raho, D. (2015). Kvm, xen and docker: A performance analysis for arm based nfv and cloud computing, 2015 IEEE 3rd Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE), pp. 1–8. http://dx.doi.org/10.1109/AIEEE.2015.7367280.
- redhat (2019). O que é virtualização? Disponível em https://www.redhat.com/pt-br/topics/ virtualization/what-is-virtualization.
- RYU Project Team (2018). RYU SDN Framework. Disponível em https://osrg.github.io/ryu-book/en/Ryubook.pdf.
- Salah, T., Zemerly, M. J., Yeun, C. Y., Al-Qutayri, M. and Al-Hammadi, Y. (2017). Performance comparison between container-based and vm-based services, 2017 20th Conference on Innovations in Clouds, Internet and Networks (ICIN), pp. 185–190. http://dx.doi.org/10.1109/ICIN. 2017.7899408.
- SHAPIRO, S. S. and WILK, M. B. (1965). An analysis of variance test for normality (complete samples)†, *Biometrika* **52**(3-4): 591–611. https://doi.org/10.1093/biomet/52.3-4.591.
- Tadesse, S. S., Malandrino, F. and Chiasserini, C. (2017). Energy consumption measurements in docker, 2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC), Vol. 2, pp. 272–273. http://dx.doi.org/10.1109/COMPSAC.2017.117.
- Turnbull, J. (2014). The docker book: Containerization is the new virtualization.
- VMware (2019). Virtualização. Disponível em https://www.vmware.com/br/solutions/virtualization.html.
- Yadav, R. R., Sousa, E. T. G. and Callou, G. R. A. (2018). Performance comparison between virtual machines and docker containers, *IEEE Latin America Transactions* **16**(8): 2282–2288. http://dx.doi.org/10.1109/TLA.2018.8528247.
- Zheng, R., Wang, H., Jin, H., Xu, D., Chen, Y., Li, X., Rao, Y. and Zhang, Z. (2020). Energy saving strategy of power system cluster based on container virtualization, 2020 Asia Energy and Electrical Engineering Symposium (AEEES), pp. 351–355. http://dx.doi.org/10.1109/AEEES48850.2020.9121491.