

IMPLANTAÇÃO DE RDC-ARQ EM AMBIENTE DOCKERIZADO: POSSIBILIDADES E DESAFIOS.

Laura Mattos da Costa¹

lauramattos.unb@gmail.com

Francisco Lopes de Sousa Júnior²

franciscolsjunior@gmail.com

Resumo

Este estudo tem como objetivo analisar a implantação de Repositórios Arquivísticos Digitais Confiáveis (RDC-Arq) em Ambiente dockerizado visto que os avanços nesta área da Ciência da Informação tem se mostrado solução eficaz para demandas acerca da gestão, preservação e acesso aos documentos de arquivo. Atualmente, os maiores desafios para as empresas e instituições são o crescente volume e diversidade de documentos produzidos e acumulados, sejam analógicos ou digitais, impactando em consideráveis custos, despesas e ações que não garantem a correta preservação dos documentos. Em contrapartida, visualiza-se os avanços no que tange aos processos de preservação, recuperação e acesso às informações armazenadas em diversos sistemas e bases de arquivo, entre os avanços, destaca-se a implantação de Repositórios Arquivísticos Digitais Confiáveis (RDC-Arq). Entende-se que o RDC- Arq é a composição de plataformas de preservação, plataformas de acesso amparados por uma sólida Política Arquivística. As plataformas mais utilizadas no Brasil são Archivemática (plataforma de preservação) e AtoM (plataforma de difusão). O software Archivemática é um conjunto de ferramentas de código aberto integradas com o objetivo de criar e processar pacotes de documentos em conformidade com o modelo funcional ISO-OAIS, que disciplina e orienta um sistema de arquivo dedicado a preservação e manutenção do acesso à informação digital por longo prazo. Já o software AtoM (Access to Memory) consiste em uma aplicação web que apresenta um modelo de plataforma que distribui a descrição arquivística de forma hierárquica, e trabalha com os padrões ISAD(G), ISAAR-CPF, ISDIAH, ISDF, e que, por sua flexibilidade, suporta outros padrões de metadados, como o RAD, o DACS, o Dublin Core e o MODS, visando acurar o acesso e difusão dos conjuntos documentais. Os autores propõem com esse estudo analisar a implementação de repositório em ambiente dockerizado, é sabido que as aplicações possuem documentação e comunidade atuante para instalação em Virtual Machine (VM), entretanto, percebe-se que várias instituições no Brasil passaram a adotar ambientes em Docker. O Docker é uma plataforma que possibilita a automatização da implantação de aplicações em ambientes isolados denominados de containers, uma solução mais rápida e funcional que proporciona o empacotamento de aplicações e suas dependências em uma unidade padrão para desenvolvimento de software, com isolamento em nível de disco, memória e processamento. Para criar um ambiente containerizado usou-se a ferramenta Docker Compose, que possibilita a criação e configuração de múltiplos containers, apresentando cada serviço que compõe a aplicação. Com a utilização do orquestrador Docker Compose viabiliza-se coordenar os processos de build e execução dos containers, atribuindo as modificações necessárias para adaptação às particularidades impostas pela infraestrutura de TI em questão. Também devem ser consideradas outras ferramentas acessórias no processo de instalação dos softwares de preservação e difusão em Docker a exemplo do servidor de automação Jenkins, e gerenciador de infraestrutura Rancher. Como resultado de pesquisa com as plataformas AtoM e Archivemática devidamente instaladas em ambiente Docker, há a integração que configura a comunicação entre os sistemas, composta por workflow onde será criado o Pacote de Submissão de Informação (SIP) que deve ser processado pelos micros serviços e softwares open-source integrados no

¹ Universidade de Brasília, Brasília/DF, Brasil.

² Universidade de Brasília, Brasília/DF, Brasil.

Archivemática, gerando então, o Pacote de Arquivamento de Informação (AIP) e Pacote de Disseminação de Informação (DIP), sendo o AIP preservado e o DIP disponibilizado através do software AtoM para a difusão e acesso cumprindo o ciclo do RDC-Arq. Por fim, verifica-se que a implantação de RDC-Arq nas instituições é um desafio possível em ambientes inovadores baseados em container há que se lembrar que independente do ambiente um Repositório Arquivístico Digital Confiável deve ser planejado, executado e auditado por meio de Políticas de TI e Arquivologia tornando o processo completo e gerenciável a longo prazo.

Palavras-chave: Repositório Arquivístico Digital Confiável, AtoM, Archivemática, Docker, Preservação.

1. Introdução

Nas últimas décadas temos vivenciado um exponencial aumento da produção de documentos em formato digital. Diante disso, os arquivistas se viram na necessidade de apropriar-se dessa nova atribuição que consiste em garantir, preservar, gerir e dar acesso a esses documentos arquivísticos digitais.

A definição de documento digital é um “documento codificado em dígitos binários, interpretável por meio de sistema computacional.” (CONARQ, 2015, p.7). Então, é importante ressaltar que “o suporte deixa de ser um dos elementos extrínsecos do documento arquivístico digital e passa a integrar seu contexto tecnológico” (RONDINELLI, 2011, p. 233).

Em meio às transformações na área arquivística, ocorre a elaboração e publicação de resoluções produzidas pela Câmara Técnica de Documentos Eletrônicos (CTDE), do Conselho Nacional de Arquivos (Conarq) visando auxiliar, orientar e padronizar requisitos tecnológicos e elaboração de políticas que perpassam esses temas nas instituições brasileiras.

Uma das grandes questões emergidas com o avanço dos documentos digitais é o debate acerca da disciplina diplomática e dos princípios arquivísticos voltados para esses documentos, dentro dessa perspectiva, é importante garantir características basilares dos documentos de arquivo que independem de seu suporte, sendo estas a fixidez, a organicidade, a naturalidade, a unicidade, a autenticidade e a imparcialidade.

Dentro desse debate, surge as contribuições do Projeto InterPARES (Pesquisa Internacional sobre Documentos Arquivísticos Permanentes em Sistemas Eletrônicos) que tem como objetivo produzir conhecimento acerca da preservação de documentos arquivísticos produzidos ou mantidos em ambiente digital. Liderado pela professora

Luciana Duranti e o professor Terry Eastwood ocorrem entre os anos de 1999 e 2001 o InterPARES 1, seguido do InterPARES 2 entre os anos de 2002 e 2007 e entre os anos de 2007 e 2012 finda o InterPARES 3. Durante todos os anos em que o projeto atuou foram imensuráveis os avanços e as contribuições para a arquivística internacional e do Brasil.

A preservação das informações digitais vai além da questão tecnológica, perpassando por questões políticas, de viabilidade econômica e cultura institucional. Ao se falar de documentos digitais, o contexto não se encerra no objeto em si e se expande para os metadados intrínsecos e extrínsecos do documento. Estes dados sobre dados auxiliam no gerenciamento, o acesso e a preservação digital a longo prazo.

O objetivo deste estudo é analisar a implantação de Repositórios Arquivísticos Digitais Confiáveis (RDC-Arq) em Ambiente dockerizado, em especial com os sistemas AtoM e Archivematica.

1.1 Contextualização do RDC-Arq

O conceito de Repositórios Arquivísticos Digitais Confiáveis (RDC-Arq) emerge no Brasil a partir da Resolução do Conselho Nacional de Arquivos (Conarq) de número 43 de 04 de setembro de 2015. Tal resolução se inspira no Modelo OAIS e na ISO 14721/2003 traduzida na norma ABNT 15471/2007 com o título “Sistema Aberto de Arquivamento de Informação- SAAI”. O objetivo explicitado na resolução do Conarq é o de:

Indicar parâmetros de forma a garantir autenticidade (identidade e integridade), a confidencialidade, disponibilidade, o acesso e a preservação, tendo em vista a perspectiva da necessidade de manutenção dos acervos documentais por longos períodos de tempo ou, até mesmo, permanentemente. (CONARQ, 2015, p.5)

Entende-se que o RDC-Arq é a composição de plataformas de preservação, plataformas de acesso amparados por uma sólida Política Arquivística. Dessa forma, os parâmetros indicados pela resolução se dividem em: Infraestrutura organizacional, apresentando quesitos como governança e viabilidade, estrutura organizacional e de pessoal, transparência de procedimentos e arcabouço político e sustentabilidade financeira. Pode-se dizer que a primeira parte visa estruturar todo o contexto necessário não só para a implantação como a manutenção e transparência do Repositórios Arquivísticos Digitais Confiável. A segunda subdivisão trata do

Gerenciamento do documento digital onde caminha desde a admissão de documentos e metadados, ao planejamento da preservação, armazenamento e manutenção, gerenciamento da informação e gerenciamento do acesso. É neste momento que é explicitado os pacotes de informação, sendo estes os documentos digitais (informação e conteúdo) bem como deus metadados (informação de representação), os pacotes são respectivamente o Pacote de Submissão de Informação (SIP), o Pacote de Arquivamento de Informação (AIP) e o Pacote de Disseminação de Informação (DIP). Na citada divisão é onde se apresentam o maior quantidade de requisitos que perpassam a configuração e estruturação do software empacotador e de micro serviços, no caso deste estudo, o Archivematica, e apresenta na parte de Gerenciamento de Acesso e com a criação do DIP a necessidade de plataforma de acesso, aqui o ICA ATOM. Por fim, é possível fazer aprofundamentos acerca das características e diplomáticas dos documentos digitais, bem como o mapeamento e caracterização dos metadados (informações de representação) dos objetos digitais do RDC-Arq. A terceira e última parte da resolução apresenta os aspectos de tecnologia, infraestrutura técnica e segurança. Apresentando boas práticas sobre gestão de dados e de segurança que devem ser garantidas pelo repositório.

1.2 O Modelo OAIS e a ISO 14721:2003

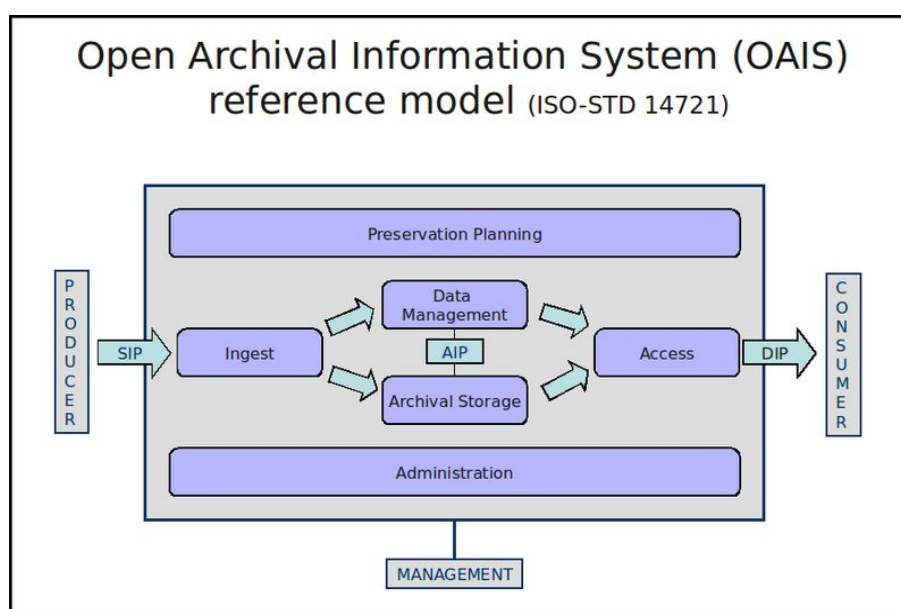


Figura 1 - Modelo de referência OAIS (Fonte: Modelo OAIS)

O modelo Open Archival Information System (OAIS), documento que inspirou a resolução brasileira, apresenta referências para um sistema de preservação e manutenção do acesso a informações digitais por longo prazo. Teve seu desenvolvimento pela Consultative Committee for Space Data System- CCSDS e tornou-se uma norma internacional, ISO 14721:2003.

O modelo divide os ambientes em: produtores (pessoas ou sistemas que fornecem a informação); Administração (responsável pelas políticas) e consumidores (pessoas ou sistemas que visam acessar as informações).

A partir do modelo, a inserção inicia com a entrada do objeto mais seus metadados, requisito obrigatório, após a captura para o sistema é feito o processamento por microserviços que ao final se forma o SIP. O objetivo é criar cápsulas dos objetos mais metadados através de hash e convertê-los para formatos de preservação através de pacotes AIP que serão preservados. Ao fim, é gerado também o DIP, o pacote de difusão que será disponibilizado para difusão e consequentemente para o consumidor.

1.3 Softwares Atom e Archivematica: Conceitualizações, papel no RDC-Arq e motivo de escolha

O software Archivematica é um conjunto de ferramentas de código aberto integradas com o objetivo de criar e processar pacotes de documentos em conformidade com o modelo funcional ISO-OAIS, que por sua vez disciplina e orienta um sistema de arquivo dedicado a preservação e manutenção do acesso à informação digital por longo prazo. Desenvolvido pela empresa canadense Artefactual Systems configura-se um software livre, gratuito e de código aberto.

A aplicação executa micro serviços, sendo estas tarefas granulares do sistema que operam sobre um modelo conceitual compatível com os pacotes definidos no modelo OAIS. Esses pacotes de informação são inseridos em um workflow e processados pelos micro serviços providos por uma combinação de scripts em Python do Archivematica e softwares open-source empacotados no sistema.

O archivematica executa importante papel no RDC-Arq e na implementação do modelo OAIS, o fluxo de submissão de pacotes começa pela aba “transfer” onde é

selecionado os arquivos a serem preservados por meio de um browser da própria aplicação que direciona ao diretório onde os arquivos deverão estar alocados, criando o Pacote de Submissão de Informação (SIP), em seguida, ocorre o “ingest” onde o SIP recebe os metadados descritivos conforme estabelecido no Plano de Preservação já cadastrado na plataforma. Após a fase de “ingest” é criado o Pacote de Arquivamento de Informação (AIP) que será destinado ao repositório que poderá ser gestado através da aba de “archival storage”, é também nessa fase em que, se escolhido, será criado o Pacote de Disseminação de Informação (DIP) que poderá ser disponibilizado para uma plataforma de difusão e acesso.

Uma vez que foi desenvolvido já visando alcançar os requisitos dos modelos consolidados, o Archivematica tem se destacado, particularmente no Brasil, por possibilitar a implantação do RDC-Arq conforme a resolução nº43/2015 da CTDE e propiciando o devido tratamento às partes lógicas e conceituais dos documentos digitais. Pelos motivos expostos, esse software foi escolhido para esse estudo de implementação em ambiente dockerizado.

A plataforma AtoM a ser apresentada a seguir é parte do Repositório Arquivístico Digital Confiável devido ao seu papel de promover e complementar a questão relacionada ao acesso às informações e objetos dos pacotes preservados. Como a resolução afirma, o RDC-Arq visa:

Indicar parâmetros para repositórios arquivísticos digitais confiáveis, de forma a garantir a autenticidade (identidade e integridade), a confidencialidade, a disponibilidade, o acesso e a preservação, tendo em vista a perspectiva da necessidade de manutenção dos acervos documentais por longos períodos de tempo ou, até mesmo, permanentemente. (Grifo nosso, CONARQ, 2015, p.5).

A parte dois da resolução tange sobre o gerenciamento do documento digital, especificamente na subdivisão de gerenciamento do acesso, é apresentado o pacote de acesso aos documentos digitais e seus metadados associados, o Pacote de informação para disseminação (DIP). É a partir desse pacote que a plataforma de difusão e acesso é integrada por meio da interoperabilidade, requisito obrigatório, ao sistema de RDC-Arq.

A plataforma escolhida para esse estudo, o software AtoM é um acrônimo para “acesso à memória”, e tem como função principal a divulgação e disponibilização de acervos arquivísticos na internet, consistindo-se em uma ferramenta de descrição

gratuita, de fácil manejo e implementação às entidades custodiadoras de todo o mundo. Baseando-se nas normas de descrição do Conselho Internacional de Arquivo (CIA), o AtoM possui uma série de normas de intercâmbio de metadados que suporta a fácil importação e exportação em padrões consolidados na comunidade como: EAD, EAC-CPF, CSV e SKOS.

O AtoM é um software fundamentado em ambiente web, sendo uma aplicação de código aberto, ou seja, pode ser usado, copiado, modificado e redistribuído sem restrições de copyright. Inicialmente construído com suporte do Conselho Internacional de Arquivos para incentivar a adoção das normas internacionais, a aplicação é desde sua versão 2.0.0 de outubro de 2013 desenvolvida e distribuída pelo grupo canadense Artefactual System, que mantém uma comunidade ativa, fator imprescindível para a constante evolução da aplicação, tanto referindo ao reporte e correção de bugs, quanto para a implementação de recursos e funcionalidades ou adaptações a novas regras e formatos que possam surgir na área arquivística.

Dessa forma, com softwares Archivematica e AtoM é possível por meio de interoperabilidade iniciar a implantação de RDC-Arq atrelado com uma Política de Preservação Institucional e um repositório digital.

A seguir, adentrarmos como se procede a instalação dos softwares escolhidos para o estudo em ambientes informatizados, analisando virtualização e containerização.

2. Metodologia de pesquisa

Buscou-se para essa pesquisa as bases da pesquisa qualitativa do tipo: Estudo de Caso Exploratório. Os problemas de pesquisa que motivaram o estudo foram: É possível instalar as plataformas que compõem o RDC-Arq em Docker? Essas plataformas interromperam? É possível garantir o fluxo correto do RDC-Arq em Docker? Quais são as possibilidades e desafios?

Dessa forma, teve-se como objetivo analisar a implantação de Repositórios Arquivísticos Digitais Confiáveis (RDC-Arq) em Ambiente dockerizado como solução eficaz para a preservação e acesso aos documentos de arquivo.

Para o estudo foram escolhidos os sistemas AtoM e Archivematica e implantados os seguintes procedimentos de coleta de dados: pesquisas bibliográficas, pesquisas na documentação dos sistemas e normas relacionadas ao tema, pesquisas em fóruns de tecnologia e testes de aplicação do modelo criado.

3. Métodos de Instalação

Neste estudo tratamos da instalação das aplicações em ambiente dockerizado em contrapartida ao ambiente de virtualização. Para tanto, é importante caracterizar cada uma das opções.

3.1 Virtualização por Máquina Virtual

A máquina virtual consiste em uma emulação de um sistema computacional e possibilita criar um cenário de várias réplicas de “máquinas físicas”, cada uma com seu sistema operacional, em uma única máquina. Em máquina virtual os sistemas operacionais e suas aplicações dividem recursos de hardware de um único servidor host ou de um conjunto de servidores hosts. Portanto, cada máquina virtual requer seu próprio sistema operacional e assim o hardware é virtualizado. Esse modelo surgiu para diminuir custos e aumentar a eficiência na TI e tem como seus benefícios a facilidade de disponibilidade de todos os recursos do sistema operacional para as aplicações e a estabilidade das ferramentas de gerenciamento e de segurança das aplicações.

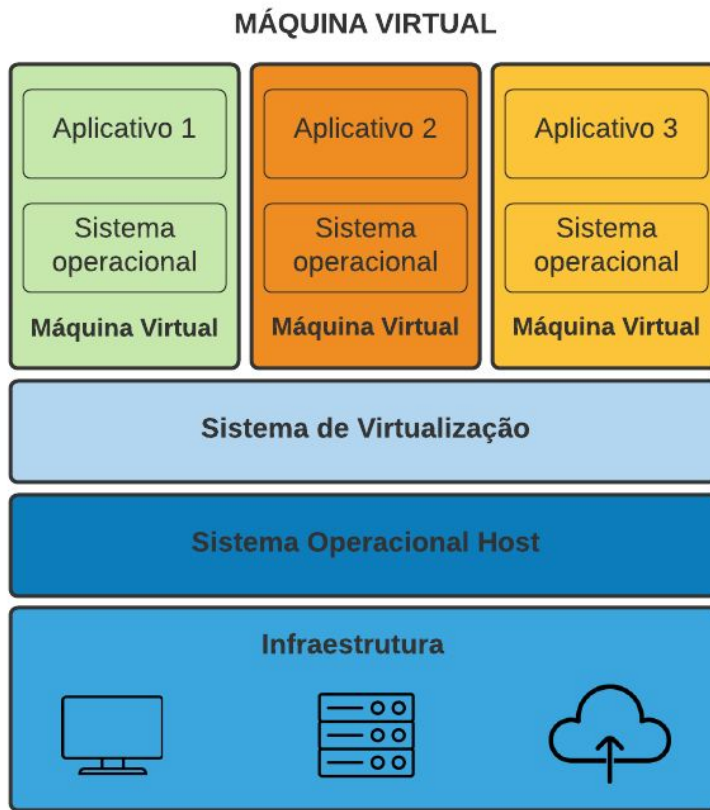


Figura 2 - Estruturação de VM. (Fonte: Elaborado pelos autores).

3.2 Virtualização por container

Como alternativa às soluções de TI por meio da virtualização completa, surgiu a tecnologia de virtualização baseada em container. A plataforma Docker permite a virtualização em nível de sistema operacional, ou seja, possibilita que vários aplicativos rodem sem executar outros kernels do sistema operacional. É possível entender os containers como máquinas virtuais modulares extremamente leves e replicáveis, facilitando a criação, economizando tempo de configuração, reduzindo a complexidade de implementação e melhorando o gerenciamento de recursos disponíveis.

O container fica no topo do servidor físico e divide o sistema operacional host e as bibliotecas. Assim, reduz significativamente a necessidade de reproduzir o código do sistema operacional o que faz com que o servidor possa rodar múltiplas cargas de trabalho com a instalação de apenas um sistema operacional.

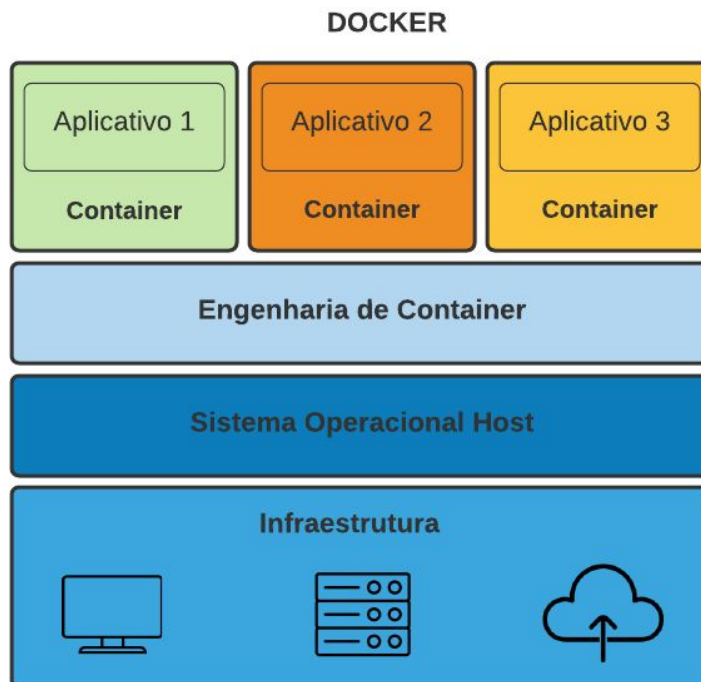


Figura 3 - Estruturação em Docker. (Fonte: Elaborado pelos autores).

Ao compararmos as características de utilização de Docker versus VM, observamos as seguintes questões:

	Docker	VM
Suporte de Sistema Operacional	Hospedado em um único servidor físico onde é possível rodar múltiplas aplicações em um único kernel.	Cada VM precisa ter seu próprio sistema operacional instalado
Segurança	Possui acesso aos subsistemas do kernel	propicia que a as aplicações acessem matriz de segurança
Portabilidade	Os contêineres possuem pacotes que podem rodar as aplicações	O servidor isolado possui o seu próprio sistema operacional

Performance	Arquitetura leve e rápida, contendo somente o necessário para atender seu serviço.	Requer maior quantidade de recursos por usar SO completo.
-------------	--	---

Quadro 1– Comparativo VM versus Docker. (Fonte: Elaborado pelos autores).

A partir das análises e testes usados como referências para este estudo, os ganhos com a utilização do Docker são perceptíveis desde os primeiros passos da implementação das aplicações AtoM e Archivematica nas instituições. A Artefactual (empresa responsável pela distribuição dos softwares) disponibiliza e indica como método de instalação padrão um manual com os passos que devem ser seguidos, contendo a instalação manual de todos os softwares e serviços ao qual as aplicações dependem, sendo eles:

AtoM	Archivematica
MySQL	Elasticsearch
Elasticsearch	Storage Service
Nginx	archivematica mcp server
Memcached	archivematica dashboard
Gearman job server	archivematica mcp client

Quadro 2 – Containers de softwares e serviços. (Fonte: Elaborado pelos autores).

Essas são algumas das dependências que exigem a instalação manual em uma única VM, quando usamos Docker cada dependência é encapsulada por um container, sendo disponibilizados e configurados em apenas um arquivo chamado `composefile.yml`.

É comum em ambientes de grande escala encontrarmos ferramentas que auxiliam uma infraestrutura com aplicações em Docker, tais como a plataforma de gerenciamento de containers Rancher, o sistema de orquestração de containers Kubernetes, Apache Mesos, Docker Swarm e serviços de integração contínua como Jenkins, Travis e GitLab CI. Essas ferramentas atuam nos processos de gerenciamento,

deploy, versionamento, escalabilidade e podem trazer uma série de benefícios que devem ser pensados previamente pela infraestrutura de TI das devidas instituições.

4. Análise de Resultados

As hipóteses de pesquisa foram consolidadas a partir de perguntas motivadoras, as primeiras foram: É possível instalar as plataformas que compõem o RDC-Arq em Docker? Essas plataformas interromperam? É possível garantir o fluxo correto do RDC-Arq em Docker? A resposta surgiu a partir da pesquisa teórica, documental e testes práticos onde obteve-se o resultado que é possível instalar as plataformas em ambiente Docker, os sistemas em suas versões 1.11.2 do Archivematica e 2.5.3 do AtoM interromperam a partir de comunicação de seus containers em um diretório comum espelhado possibilitando o ciclo de inserção de pacotes arquivísticos no repositório digital e devidamente representado na aplicação de acesso ao público através do envio do DIP à plataforma de difusão. Assim, constatou-se que o fluxo funcionou corretamente de igual forma a experienciado em ambiente de máquina virtual.

Questionamos ainda quais são as possibilidades e desafios da Implantação do RDC-Arq em ambiente dockerizado. Os resultados obtidos acerca das possibilidades apontaram três grandes vantagens, sendo elas: A replicabilidade a partir da ferramenta Docker compose que possibilita a implementação de sistemas conhecidos em arquivos (Atom, Archivematica, Alfresco, outros) a partir de estrutura de arquitetura pré definida trazendo eficiência a instalação e configuração dessas ferramentas e seus fluxos de trabalho em unidades de arquivo. A implementação de configurações semelhantes em diferentes órgãos propicia a gestão de um RDC-Arq composto por fundos de múltiplas instituições, a exemplo de possível aplicação por um Arquivo Nacional que é custodiante de diferentes fundos. A replicabilidade possibilita ainda testes isolados propiciando mais segurança e liberdade aos desenvolvedores. Outra importante constatação foi a implementação de serviços de gestão de pipelines e facilidade na integração contínua ocasionando em agilidade em corrigir erros e realizar novos deploys. A última vantagem destacada é a promoção do uso eficiente dos recursos de sistemas onde as instâncias dos aplicativos usam muito menos memória

em ambiente Docker do que as VMs, caracterizado por inicializações e interrupções rápidas resultando em drástica redução de gastos em recursos de TI.

Os resultados do trabalho apontaram relevantes contribuições teóricas ao abordar e introduzir o estudo de uma nova ferramenta da TI para uso da implementação de premissas arquivísticas, onde a ferramenta Docker possui baixíssima ou nenhuma citação em artigos científicos da área da Arquivologia e da Ciência da Informação, colaborando pro desenvolvimento dos temas relacionados à TI nessas disciplinas. No campo prático a utilização do Docker para implementação de RDC-Arq se mostrou possível e responsável por uma série de benefícios e possibilidades que envolvem desde a facilidade na instalação e configuração por profissionais de fora da TI até o advento da replicabilidade que pode ser utilizadas por unidades de arquivo custodiadora de múltiplos fundos, abarcando ainda os facilitadores em relação aos recursos de TI.

O estudo apresentou desafios que podem ser considerados limitações até que sejam melhor exploradas, entre elas destacamos que Artefactual (empresa responsável pelos softwares Atom e Archivemática) não utiliza Docker como ambiente de produção. Assim, podemos encontrar bugs ou inconsistências em atualizações que não foram pensadas para Docker. Outro fator é a escassez de material de referência na documentação dos sistemas e publicações em fóruns de debate. Um importante desafio observado é a necessidade do profissional responsável pela instalação em lidar com diferentes sistemas relacionados ao Docker: Jenkins, Rancher, Kubernetes, Docker Swarm, entre outros. Por fim, há de se ressaltar a necessidade de estudar e planejar alterações e manutenção da segurança e de atualizações a longo prazo do RDC-Arq em ambiente Docker. Todos os resultados e suas implicações teóricas e práticas apresentam considerável grau de confiabilidade uma vez que foram testadas sob o formato de estudo de caso.

Como expectativas de pesquisa futuras estão a observação e análise do uso do Docker nas plataformas Atom e Archivemática ao longo de novas versões e a manutenção e governança acerca das manutenções e segurança. Se deseja ainda buscar interoperabilidade entre o RDC-Arq, SIGAD e sistemas de negócios como forma de garantir a cadeia de custódia ininterrupta em ambiente dockerizado.

5. Considerações Finais

O objetivo deste trabalho foi analisar a implantação de Repositórios Arquivísticos Digitais Confiáveis (RDC-Arq) em Ambiente dockerizado como solução eficaz para a preservação e acesso aos documentos de arquivo verificando sua condições de instalação, configuração, interoperabilidade, adequação aos parâmetros e possibilidades e desafios. A motivação se pautou nos avanços nesta área da Ciência da Informação que tem se mostrado solução eficaz para demandas acerca da gestão, preservação e acesso aos documentos de arquivo, entretanto, percebemos que há argo conhecimento sobre a instalação em formato de máquina virtual e pouco conteúdo e exemplos sobre a inovação da implantação em Docker.

O primeiro passo do trabalho foi analisar o tema através de pesquisas bibliográficas nas áreas de Arquivologia e Ciência da Informação, pesquisas na documentação dos sistemas e normas relacionadas ao tema sobre a instalação e utilização dos sistemas escolhidos em ambiente Docker e pesquisas em fóruns de discussão. Entretanto, majoritariamente as pesquisas não trouxeram resultados relevantes e reforçaram a afirmação que praticamente não há estudos sobre esse tema.

Paralelamente, foi iniciada uma parte prática de instalação e registro da implementação dos softwares em ambiente dockerizado. Nessa parte foi moldado o projeto de cada uma das aplicações para que melhor se encaixasse com a infraestrutura do órgão alvo da pesquisa. Tendo alterações nos projetos para publicação no repositório interno, e adequações para o ambiente de cluster junto às ferramentas dispostas para a gestão dos mesmos, como Rancher e Jenkins.

A última parte do trabalho foi a consolidação do estudo, registrando os resultados e identificando as possibilidades, desafios e estudos futuros.

A conclusão final é que através do estudo desenvolvidos afirmamos que a utilização de Docker em sistemas informatizados de gestão de documentos e Repositórios Arquivísticos Digitais Confiáveis (RDC-Arq) ainda é um grande campo a ser explorado, mas que em, estudos iniciais se mostra muito promissor.

Referências

ARQUIVO NACIONAL (Brasil). *Dicionário brasileiro de terminologia arquivística*. Rio de Janeiro: Arquivo Nacional, 2005. 232 p. (Publicações Técnicas; n 51). Disponível em: <<http://simagestao.com.br/wp-content/uploads/2016/01/Dicionario-de-terminologia-arquivistica.pdf>>. Acesso em: set. 2020.

BRASIL. Arquivo Nacional. Conselho Nacional de Arquivos (2015). *Diretrizes para a implementação de repositórios arquivísticos digitais confiáveis– RDC-Arq*. Rio de Janeiro, 2015. Disponível em: <http://www.conarq.gov.br/images/publicacoes_textos/diretrizes_rdc_arq.pdf>. Acesso em: set. 2020.

BRASIL. Conselho Nacional de Arquivos. Câmara Técnica de Documentos Eletrônicos. *Glossário Documentos Arquivísticos Digitais*. Rio de Janeiro: CONARQ, 2020. Disponível em: <http://antigo.conarq.gov.br/images/ctde/Glossario/glosctde_2020_08_07.pdf>. Acesso em: set. 2020.

ISO. International Organization for Standardization. *ISO 15489-1: Information and documentation -- Records management - Part 1: General*. Geneva/Switzerland : ISO, 2001.

ISO. International Organization for Standardization. *ISO 23081-1: Information and documentation -- Records management - Part 1: General*. Geneva/Switzerland : ISO, 2006.

BRASIL. Lei nº 8.159, de 8 de janeiro de 1991. *Dispõe sobre a política nacional de arquivos públicos e privados e dá outras providências*. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/L8159.html>. Acesso em: set. 2020.

RONDINELLI, Rosely Curi. *O Conceito de documento arquivístico frente à realidade digital: uma revisão necessária*. Tese (Doutorado em Ciência da Informação) –Universidade Federal Fluminense, PPGCI. Niterói, 2011.

GRÁCIO, José Carlos Abbud. *Preservação digital na gestão da informação: um modelo processual para as instituições de ensino superior*. São Paulo: Cultura Acadêmica, 2012. Disponível em: <<http://hdl.handle.net/11449/113727>>. Acesso em: set. 2020.