

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS EXATAS
DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO
ESPECIALIZAÇÃO EM INFORMÁTICA: ÁREA DE CONCENTRAÇÃO: GESTÃO DE
TECNOLOGIA DA INFORMAÇÃO

Danilo Almeida Paiva

**VIRTUALIZAÇÃO DA INFRAESTRUTURA: SISTEMATIZAÇÃO E PESQUISAS
CORRENTES**

Brasília

2018

DANILO ALMEIDA PAIVA

**VIRTUALIZAÇÃO DA INFRAESTRUTURA:
SISTEMATIZAÇÃO E PESQUISAS CORRENTES**

Monografia apresentada ao Curso de Especialização em Informática do Departamento de Ciência da Computação do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais, como requisito parcial para a obtenção do grau de Especialista em Informática.

Área de Concentração: Gestão de Tecnologia da Informação

Orientador: Dorgival Olavo Guedes Neto

Brasília
2018

© Danilo Almeida Paiva, 2019
Todos os direitos reservados

Ficha catalográfica elaborada pela Biblioteca do ICEx – UFMG

Paiva, Danilo Almeida

P149v Virtualização da infraestrutura: sistematização e pesquisas
correntes / Danilo Almeida Paiva – Brasília, 2019.
x, 40 f., il.

Monografia (especialização) – Universidade Federal de
Minas Gerais. Departamento de Ciência da Computação.

Orientador: Dorgival Olavo Guedes Neto

1.Computação – Monografias. 2. Virtualização de
infraestrutura. 3. Data center. 4. virtualização de armazenamento.
5. Administração Pública. I. Orientador. II. Título

CDU 519.6*



UNIVERSIDADE FEDERAL DE MINAS GERAIS

INSTITUTO DE CIÊNCIAS EXATAS
DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO
ESPECIALIZAÇÃO EM INFORMÁTICA: ÁREA DE CONCENTRAÇÃO GESTÃO EM
TECNOLOGIAS DA INFORMAÇÃO

Virtualização: Sistematização e Pesquisas Correntes

DANILO ALMEIDA PAIVA

Monografia apresentada aos Senhores:

Prof. Dergival Olavo Guedes Neto
Orientador
DCC - ICEX - UFMG

Prof. José Nagib Cotrim Árabe
DCC - ICEX - UFMG

Prof. José Marcos Silva Nogueira
DCC - ICEX - UFMG

Belo Horizonte, 14 de março de 2019

Este trabalho é dedicado à Diogo da Fonseca Tabalipa por, enquanto gestor, compreender a necessidade de atualização dos servidores públicos de TIC sob sua coordenação, incentivando-os à obtenção de conhecimento.

Aos colegas de pós-graduação e aos professores, pelo caminho que trilhamos juntos.

Agradecimentos

A conclusão deste trabalho não seria possível sem o apoio da colega Adriana Barbosa Lima, que, através de aplicativo de troca de mensagens (whatsapp), sugeria tópicos, acrescentava idéias importantes ao tema desta Monografia e direcionava a forma de abordagem.

Agradeço também a outro colega: Romário César de Almeida, pela troca de informações na execução dos trabalhos do curso de especialização.

RESUMO

Este trabalho versa sobre a Virtualização de Infraestrutura de Data center (Centro de Processamento de Dados), mais especificamente sobre a virtualização de servidores, de armazenamento e de redes. Esses tópicos são considerados importantes na otimização do uso de recursos de TIC considerando a maturidade tecnológica da Administração Pública Federal (APF). O trabalho foca em base histórica e principais conceitos dentro dos contextos de virtualização elencados para facilitar o diálogo entre servidores públicos de TIC da APF e fornecedores na busca por soluções operacionais relacionadas às necessidades estratégicas organizacionais.

Palavras-chave: virtualização de infraestrutura; data center; virtualização de servidores; virtualização de armazenamento; virtualização de redes; Administração Pública Federal.

ABSTRACT

This work is about the Virtualization of Data Center Infrastructure, specifically about server, storage and network virtualization. These topics are considered important in the optimization of the use of ICT resources considering the technological maturity of the Federal Public Administration (APF). The work focuses on historical basis and key concepts within the virtualized contexts listed to facilitate the dialogue between APF's public ICT servants and suppliers in search of operational solutions related to strategic organizational needs.

Keywords: infrastructure virtualization; data center; server virtualization; storage virtualization; network virtualization; Federal Public Administration.

Lista de Siglas

ATIs	Analistas em Tecnologia da Informação
CGTIs	Coordenações-Gerais de Tecnologia da Informação
CTIC	Data Center do Centro de Tecnologia da Informação e Comunicação
DAS	Direct Attached Storage
DTI	Diretoria de Tecnologia da Informação
DTIs	Diretorias de Tecnologia da Informação
FC	Fibre Channel
FCoE	Fibre Channel over Ethernet
IBRAM	Instituto Brasileiro de Museus
MCT	Ministério da Ciência, Tecnologia, Inovação
MCTIC	Ministério da Ciência, Tecnologia, Inovações e Comunicações
MPDG	Ministério do Planejamento, Desenvolvimento e Gestão
SDDC	Software Defined Datacenter
SO	Sistema Operacional
TCO	Custo Total de Propriedade
TCOs	custos totais de propriedade
TIC	Tecnologia da Informação e Comunicação
HDD	hard-disk drives
NAS	Network Attached Storage
NFV -	Network Function Virtualization
USP	Universidade de São Paulo
RAID	Redundant Array of Independent Disks;
SAN	Storage Area Network
SDN	Software Defined Networking
VM	máquina virtual

Lista de ilustrações

Figura 1 – Planejamento de contratação de solução de armazenamento ativo-ativo em março de 2018 para o MCTIC considerando a distância entre edifícios (Blocos E e R da Esplanada dos Ministérios), a latência da rede e discos all flash.	11
Figura 2 – Projeto de replicação de dados e estrutura de virtualização necessária considerando a tecnologia do fabricante VMware.	12
Figura 3 – Evolução do Data Center - 1.0, 2.0 e 3.0 conforme (SANTANA, 2013)	15
Figura 4 – Virtualização do Data Center com escopo deste trabalho - caminho para a utilização de nuvem	15
Figura 5 – Modelo tradicional (uma aplicação/componente por servidor) versus Virtualização de Servidores (n aplicações/componentes para 1 servidor)	17
Figura 6 – Tipos de Hypervisors 1(baremetal) e 2 (hosted)	18
Figura 7 – Virtualização Total x Paravirtualização	19
Figura 8 – Quadrante-mágico do Gartner para soluções de virtualização de servidores x86	20
Figura 9 – Cálculo do licenciamento Microsoft anual, com base na infraestrutura citada. A ferramenta online disponível no site http://datacenter-tco-tool.azurewebsites.net/ ainda compara com o preço anual da solução VMware.	23
Figura 10 – Virtualização tradicional com o uso do VMware vSphere versus o uso de Kubernetes como orquestrador de contêineres.	26
Figura 11 – Discos Mecânicos vs SSD	28
Figura 12 – DAS, NAS e SAN - Tipos de Redes de Armazenamento	30
Figura 13 – Rede Tradicional vs SDN	32
Figura 14 – Aplicações, Camada de Controle (SDN Controller), Camada de Dados (switch). A comunicação entre SDN Controller e os vswitches ocorre por meio do protocolo Openflow (forma mais comum).	32
Figura 15 – emphAppliances de rede vs NFVs	33

Sumário

1	.	
1	Introdução	10
2	Revisão da Literatura	13
2.1	Virtualização: Conceito amplo	13
2.2	Evolução do Data center	14
2.3	A origem da virtualização	15
2.4	Virtualização de Servidores	16
2.4.1	Hypervisor ou Máquina Virtual de Sistema	17
2.4.2	Tipos de Virtualização	18
2.4.3	Conceito de workload e throughput	19
2.4.4	Fornecedores conforme Gartner	19
2.4.5	Vantagens da Virtualização de Servidores	20
2.4.6	Desvantagens da Virtualização de Servidores	22
2.4.7	Considerações recentes . .	24
2.4.7.1	Virtualização e a utilização de Contêineres	25
2.5	Virtualização de Armazenamento	26
2.5.1	O Sistema de Armazenamento	27
2.5.2	Os tipos de Storages	28
2.5.3	Tipos de Rede de Armazenamento	29
2.5.4	Virtualização	30
2.5.5	Storage Definido por Software (SDS)	30
2.6	Virtualização de Redes	30
2.6.1	Conceito	30
2.6.2	Paradigmas	31
2.6.2.1	SDN	31
2.6.2.2	NFVs (Funções Virtualizadas de Redes)	32
2.6.3	Considerações recentes	33
3	Materiais e Métodos	34
4	Resultados	35
5	Conclusão	36
	Bibliografia	37

1 Introdução

Esta pesquisa versa sobre o tema virtualização de recursos de infraestrutura de TIC, com foco na Administração Pública Federal. Conhecer suas motivações, faz-se imprescindível.

Os Analistas em Tecnologia da Informação (ATIs) são profissionais concursados (nível superior) da Esfera Federal do Executivo para exercerem funções de planejamento, supervisão, coordenação e controle dos recursos de tecnologia da informação (MPDG, 2018a). Eles atuam nos órgãos que fazem parte do SISP (Sistema de Administração dos Recursos de Tecnologia da Informação), dessa forma podendo estar descentralizados em vários órgãos (já que são oriundos do antigo Ministério do Planejamento, Desenvolvimento e Gestão - atual Ministério da Economia). Há a mobilidade dos ATIs que podem de tempos em tempos trocar de órgão - respeitando-se alguns critérios. Grande parte dos ATIs estão no Ministério da Economia. O restante está alocado dentro das Diretorias de Tecnologia da Informação (DTIs) ou Coordenações-Gerais de Tecnologia da Informação (CGTIs). Outros cargos de TIC (Tecnologia da Informação e Comunicação) da Administração Pública Federal estão geralmente fixados aos órgãos que criaram os respectivos concursos para provimento de funções específicas.

As DTIs e CGTIs são geralmente divididas de forma estratégica em setores comuns para que sejam realizadas ações coesas e alinhadas com as diretrizes do governo federal. São eles: (1) Governança; (2) Sistemas; (3) Infraestrutura. Os servidores públicos de TIC podem mudar de atribuições condizentes com esses três setores, mas é necessário um aprendizado teórico ou tecnológico para exercerem as possíveis atribuições. A função do servidor público de TIC tende a ser mais mutável do que na iniciativa privada em que se constrói um currículo ou vida profissional direcionados para funções específicas.

Tendo em vista essa variação de atribuições e buscando-se uma facilidade de aprendizado, no que se refere à funções no setor de Infraestrutura de TIC, também na subárea de Aquisições inserida geralmente dentro da Governança, **entender os conceitos sobre virtualização é de extrema importância.**

Esta pesquisa tem foco nos servidores públicos de TIC, consolidando informações técnicas sobre esse tema que são imprescindíveis no contexto operacional da transformação digital dos serviços públicos pretendida pelo Governo Federal. Em relatório de transição de governo elaborado pelo então Ministério do Planejamento, Desenvolvimento e Gestão (MPDG, 2018b), enfatiza-se esse caminho:

A transformação digital no governo é uma oportunidade para tornar o Estado

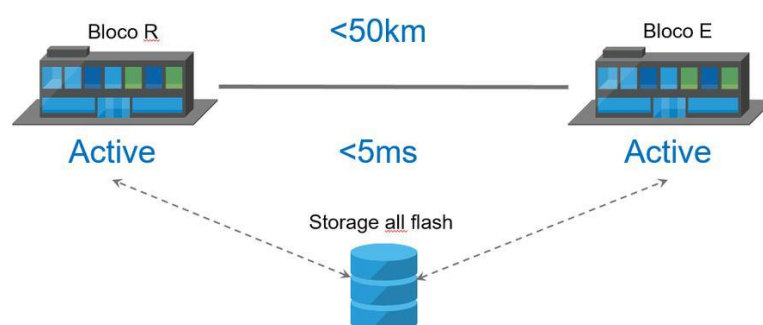
mais eficiente, dado que as tecnologias facilitam a alocação adequada de recursos humanos e materiais, permitindo resultados efetivos das políticas públicas.

Compreende-se a camada de virtualização nesse contexto como parte da infraestrutura de TIC necessária para provimento dos serviços digitais. Portanto, visa-se com este trabalho consolidar conceitos geralmente espalhados em literaturas diversas sobre a virtualização e pesquisas recentes buscando-se facilidade de compreensão sobre o tema.

A pesquisa em questão nasceu de uma necessidade relacionada às ações de governo. Em maio de 2016 foi publicada a Medida Provisória (MP) nº 726, convertida na Lei nº 13.341, de 29 de setembro de 2016. A lei extinguiu o Ministério das Comunicações (MC) e transformou o Ministério da Ciência, Tecnologia, Inovação (MCTI) em Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC) (BRASIL.PRESIDÊNCIA DA REPÚBLICA, 2016).

A questão que a fusão desses dois Ministérios, localizados fisicamente em edifícios distintos (Blocos E e R) da Esplanada dos Ministérios em Brasília - um quase de frente para o outro - exigiu que a TI pensasse em formas para se administrar dois data centers independentes. Ações foram feitas em direção à virtualização dos data centers e a utilização da estrutura total como dois sites (no sentido de estruturas completas de infraestrutura de TIC) para, em caso de problemas em um site, o outro fosse capaz de funcionar corretamente - redundante - mantendo todos os usuários de TIC do órgão produtivos e garantindo-se alta disponibilidade.

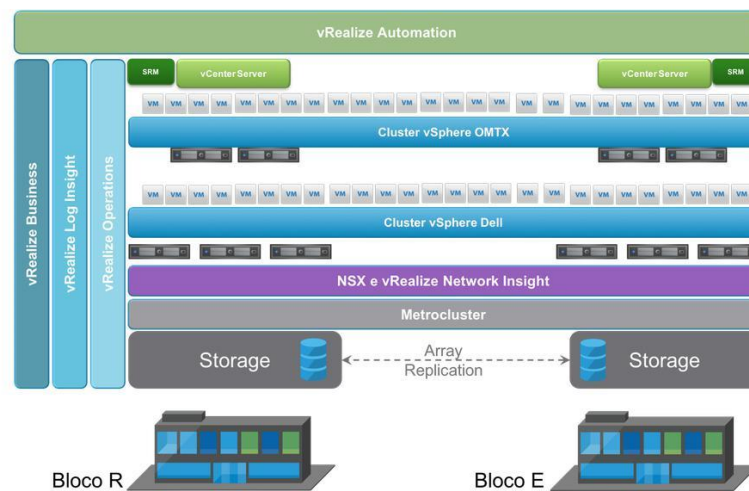
Figura 1 – Planejamento de contratação de solução de armazenamento ativo-ativo em março de 2018 para o MCTIC considerando a distância entre edifícios (Blocos E e R da Esplanada dos Ministérios), a latência da rede e discos all flash



A conversa sobre replicação de dados, conforme **FIGURA 1** provocou toda a descoberta pelos requisitos de virtualização necessários para o funcionamento da estrutura. O extinto Ministério das Comunicações incorporado ao MCTIC já possuía licenças do fabricante VMware para virtualização de servidores e um pouco depois da fusão efetivou-se uma contratação que já estava em andamento de licenças que permitiam a criação de nuvem privada (também da VMware) e funcionalidades mais

avançadas para provisionamento de recursos virtuais. O projeto final foi consolidado na **FIGURA 2**, porém não completamente foi implementado por divergência de prioridade de ações internamente na Diretoria de Tecnologia da Informação (DTI) do órgão.

Figura 2 – Projeto de replicação de dados e estrutura de virtualização necessária considerando a tecnologia do fabricante VMware



Em outro órgão, no **Instituto Brasileiro de Museus (IBRAM)**, autarquia vinculada ao antigo Ministério da Cultura e atual Ministério da Cidadania há a constante preocupação com o Data Center da instituição: ele está abrigado em uma sala imprópria dentro do edifício onde a instituição está localizada (Setor Bancário Norte em Brasília). Isso porque os requisitos de energia elétrica estável e refrigeração (muito importantes dentre outros) estão sendo parcialmente atendidos, o que gera riscos.

Além das condições supracitadas, os equipamentos de infraestrutura de TIC do IBRAM estão em sua maioria sem garantia e suporte técnico. Nesse cenário, já foram conduzidas ações para virtualizar os servidores físicos e migrar todas as aplicações para esse ambiente por questões de continuidade de negócio em caso de falhas, otimização de energia, maior disponibilidade dos serviços. O diálogo sobre virtualização permite possibilidades de ação atuais e futuras e essa é a provável realidade de vários órgãos públicos em diferentes graus de maturidade organizacional.

Pretende-se, para facilitar o diálogo sobre o tema em questão, estabelecer **cronologicamente o advento dos tipos diferentes de virtualização e os conceitos relacionados**. Desta forma, as seções conceituais descritas no TÓPICO 2 - Revisão de Literatura seguem uma ordem temporal. Os conceitos são acrescidos de informações existentes em artigos recentes.

2 Revisão da Literatura

2.1 Virtualização: Conceito amplo

(WEILL; ROSS, 2010) enfatizam que a **infraestrutura de TIC é o alicerce do modelo operacional das organizações baseadas em informação e esse modelo define a estratégia das organizações**, sendo também responsável pela possibilidade de inovação. No contexto deste trabalho enfatizam-se dois componentes grandes da infraestrutura de TIC: 1) o Data Center - o conjunto de equipamentos físicos, conexões e de edificação; 2) a Virtualização - a camada lógica.

Em 2012, 40% dos *CIOs* (*Chief Information Officers* – Diretores de TI) citaram a virtualização como tópico prioritário (considerando-se tecnologias existentes em um data center ou centro de processamento de dados) (SANTANA, 2013). Como há contínuo aumento do conjunto de tecnologias de virtualização, acredita-se que elas deixarão de ser elementos opcionais em breve, se tornando ubíquas.

Emular é sinônimo de “imitar” comportamento; abstrair, no contexto da computação, significa criar uma representação equivalente que mantenha características e propriedades em relação ao elemento que está sendo observado. Assim dito, o conceito de virtualização está relacionado ao de emulação e também ao de abstração uma vez que:

Virtualização é a emulação transparente de um recurso de TIC produzindo aos seus usuários (consumidores) benefícios que não estavam disponíveis na forma física do recurso (SANTANA, 2013). A camada de software que emula o componente físico é considerada uma abstração. Outra definição pertinente:

Virtual significa “não fisicamente existente” mas feito por software para parecer como se fosse”. Assim sendo, dentro de um contexto, pode-se dizer que um elemento virtual é uma abstração particular do elemento em si.

De uma forma geral, os benefícios com o uso de virtualização estão relacionados com: 1) redução de custo e maior utilização dos ativos; 2) maior disponibilidade e estabilidade dos recursos; 3) simplificação de processos de trabalho; 4) maior flexibilidade para inovar, ao permitir maior adaptabilidade à infraestrutura de TIC.

Não menos importante é conhecer e contornar as limitações trazidas pela virtualização. São exemplos: 1) há certa complexidade no gerenciamento\conhecimento do licenciamento existente de diversos fabricantes; 2) há a falta de profissionais especializados.

Visando-se a otimização de recursos em um data center popularizou-se a virtualização de servidores e vem se discutindo muito a de armazenamento e a de redes

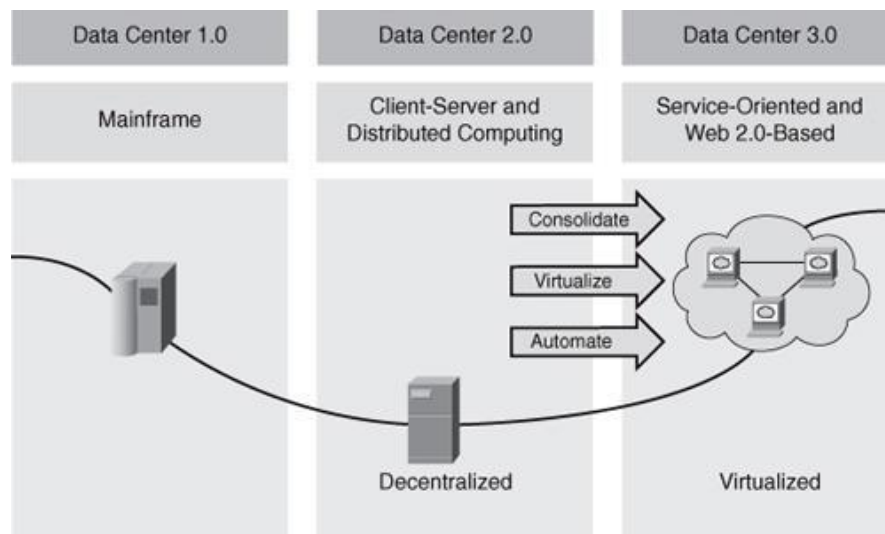
de forma mais ampla e estabelecidos conceitos relacionados dentro de cada tipo de segmento.

No atual ano de 2019 emprega-se o termo de marketing: Software Defined Datacenter (SDDC) como sendo a extensão de conceitos de virtualização tais como abstração, pooling (agregação de recursos) e automação para todos os recursos e serviços do Data center para se atingir a IT como serviço (ITaaS) (WIKIPEDIA CONTRIBUTORS, 2018). **A utilização do termo virtualização neste momento, portanto, está relacionada com a utilização de recursos de infraestrutura interdependentes existentes em um data center.**

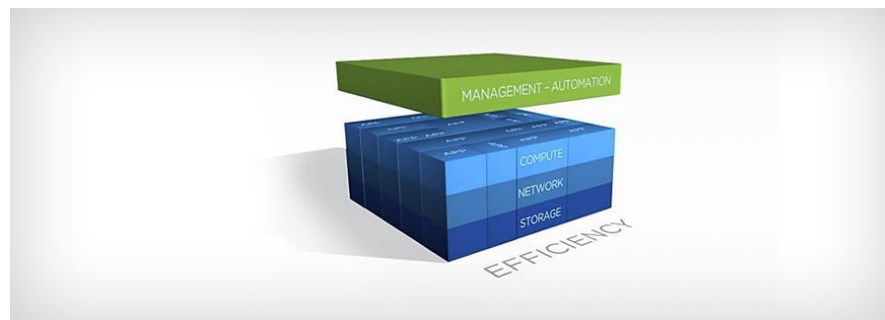
2.2 Evolução do Data center

Data center pode ser entendido como o conjunto de recursos de infraestrutura de Tecnologia da Informação e Comunicação abrigados em uma estrutura complexa e geralmente especial (com estruturas específicas de construção) em que são considerados: os tipos diferentes de equipamentos, a sua organização e disposição em racks, o consumo de energia, a necessidade de refrigeração, o cabeamento estruturado, sistemas de segurança e monitoramento; o seu gerenciamento é feito por equipe especializada (SANTANA, 2013).

SANTANA (2013), define três fases da evolução dos data centers, conforme FIGURA 3: **(1)** a partir 1950, salas especialmente construídas para abrigar mainframes com sistemas centralizados baseados em arquitetura de software monolítica (ex. IBM); **(2)** a partir de 1980 com o modelo cliente-servidor e o barateamento de servidores (arquitetura x86), os recursos de TIC ficavam mais próximos dos computadores que os acessavam, portanto os data centers eram distribuídos e improvisados - podia-se dizer que não era mais só uma sala centralizada com os recursos de TIC; **(3)** - a atual - a partir de 2000: com o avanço da interconexão das redes, os data centers novamente foram consolidados em uma estrutura única por questões de custo, segurança e administração, com servidores (equipamentos) executando funções específicas consonantes com o modelo de aplicação em camadas, porém subutilizados. Os custos para expansão dessa estrutura e questões relacionadas à energia vieram à tona promovendo estudos para a melhor utilização dos recursos existentes no Data Center (servidores, armazenamento, redes), simplificação das operações de gerenciamento dos ativos: um dos principais caminhos foi o da virtualização e mais recentemente o da automação dos recursos virtualizados/utilização de infraestrutura de terceiros (nuvem) **(FIGURA 4).**

Figura 3 – Evolução do Data Center - 1.0, 2.0 e 3.0 conforme (SANTANA, 2013)

Fonte: <http://www.ciscopress.com/articles/article.asp?p=1717681&seqNum=2>

Figura 4 – Virtualização do Data Center com escopo deste trabalho - caminho para a utilização de nuvem

Fonte: <https://www.vmware.com/products/datacenter-virtualization.html>

2.3 A origem da virtualização

A origem da virtualização (de equipamento) remonta-se à publicação do artigo Time Sharing Processing in Large Fast Computers de Christopher Strachey em **1959**. A publicação visava à melhor utilização de máquinas de grande porte através da multiprogramação em tempo compartilhado (SOUSA NETO, 2016). Entende-se como tempo compartilhado:

Os tempos ociosos entre processos são compartilhados com outros processos para dinamizar o sistema. Múltiplas tarefas são executadas simultaneamente, sabendo que a CPU despende um tempo para cada tarefa sequencialmente. Os tempos dedicados com cada tarefa são pequenos o suficiente para que os usuários interajam com cada programa sem perceber que outros estão em execução (BUENO, 2009).

Remonta-se também a esse período o desenvolvimento da “**Memória virtual**”, a expansão da memória principal dos computadores através do uso da memória secundária. Aliada ao multiprocessamento, era possível então a especificação de partições ou espaços de endereçamento que eram utilizados por programas diferentes (SOUSA NETO, 2016). Essas tecnologias foram disponibilizadas na família de equipamentos IBM /370 na década de 70.

A virtualização como entendemos hoje (abstração de plataforma física) inicialmente foi implementada como emulação de sistemas operacionais, as denominadas **máquinas virtuais de processo**. Ela surgiu para viabilizar que uma aplicação desenvolvida para um Sistema Operacional (SO) “A” funcionasse em um SO “B” que emulasse o comportamento do “A”, desta forma: binários de um processador eram interpretados e substituídos por outro processador. Essa técnica, contudo, não é eficaz pois não aproveita plenamente o hardware físico (existe considerável perda de desempenho na tradução dos códigos). Além disso, essa forma de virtualização oferece, geralmente, dispositivos genéricos de leitura/escrita que sobrecarregam o SO. **O hypervisor, ou camada de virtualização entre o hardware e o sistema operacional foi a solução que melhorou (em relação às máquinas virtuais de processo) a questão de desempenho dos dispositivos oferecidos de leitura e escrita além de possibilitar o isolamento das diferentes máquinas virtuais criadas na plataforma.**

A partir do início da década de 80 houve a popularização da arquitetura x86 nos computadores e servidores. Na década de 90, com a disponibilidade de redes de longa distância, a larga adoção do Windows e Linux como sistemas operacionais, a possibilidade do uso da virtualização se tornou objeto para a otimização do uso dos equipamentos e redução do TCO (Custo Total de Propriedade) do Data Center. Em 1998 a VMware desenvolveu o primeiro hypervisor para plataformas x86, quando conseguiu contornar a problemática das instruções privilegiadas/não privilegiadas no processador.

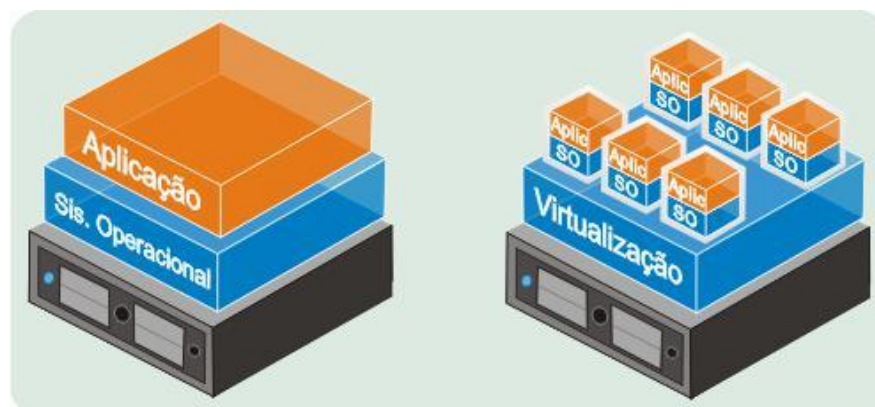
2.4 Virtualização de Servidores

A partir de 1998, as tecnologias de virtualização de servidores evoluíram e atualmente são amplamente utilizadas. A maior parte dos órgãos federais já utiliza a virtualização de servidores por ser tecnologia consolidada. O conceito principal envolvido na virtualização de servidores é **o particionamento de um equipamento servidor físico em vários servidores lógicos**. Desta forma transforma-se a forma como são hospedadas as aplicações: do padrão de uma aplicação ou componente de sistema por equipamento (1:1), passando-se a ter várias máquinas virtuais (entidades isoladas umas das outras) que executam em um mesmo equipamento (n:1), conforme **FIGURA 5**. A consolidação de servidores, o fenômeno de se administrar um

menor conjunto de servidores geralmente em cluster (agrupamento) virtualizado é o que permite: 1) a redução do espaço físico utilizado pela infraestrutura de TIC; 2) a otimização do uso dos equipamentos com a distribuição da carga de trabalho; 3) a melhor recuperação de desastres e continuidade de negócio.

Em meados da década de 2000, fabricantes de processadores otimizaram a virtualização em seus produtos (são exemplos: AMD, Intel). Foram incluídas funcionalidades específicas no hardware visando melhor desempenho das aplicações (DEVMEDIA, 2014).

Figura 5 – Modelo tradicional (uma aplicação/componente por servidor) versus Virtualização de Servidores (n aplicações/componentes para 1 servidor)



(DEVMEDIA, 2014)

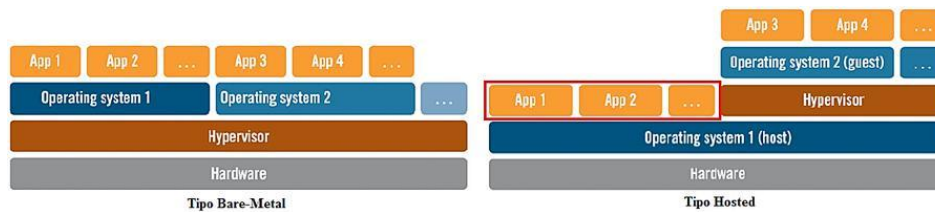
2.4.1 Hypervisor ou Máquina Virtual de Sistema

O **Hypervisor** (termo cunhado nas primeiras implementações de virtualização em máquinas IBM) ou **Máquina Virtual de Sistema** é a camada de software que é responsável pela criação de máquinas virtuais isoladas (VMs) que compartilham os mesmos recursos físicos. São responsáveis pela gerência de memória, manutenção do estado das máquinas virtuais, administrar o uso dos recursos físicos virtualizados, e é através dessa camada que se pode alocar dinamicamente mais recursos de hardware. São habitualmente classificados em dois tipos: **tipo 1** e **tipo 2**.

O primeiro tipo de Hypervisor é o classificado como **baremetal** e executa diretamente sobre o hardware físico. As máquinas virtuais (com sistemas operacionais visitantes) estão situadas acima dele. Essa implementação corresponde (teoricamente) ao que foi implementado nos sistemas IBM na década de 70. São exemplos de ferramentas que funcionam dessa forma: **VMware ESX Server**, **Microsoft Hyper-V** e **Citrix Xen Server** (SOUSA NETO, 2016). Do **tipo 1**, ainda existem subclassificações: (1) hypervisor monolítico: emula todo os drivers (de dispositivos) para as VMs, pois eles estão no hypervisor ; (2) hypervisor microkernelizado: os drivers estão dentro das VMs.

O segundo tipo corresponde ao **hosted** em hypervisor é uma aplicação que funciona como ambiente para outras aplicações. Por ser menos performático é utilizado em laboratórios, para estudo ou para testes (ex.: **VirtualBox**).

Figura 6 – Tipos de Hypervisors 1(baremetal) e 2 (hosted)



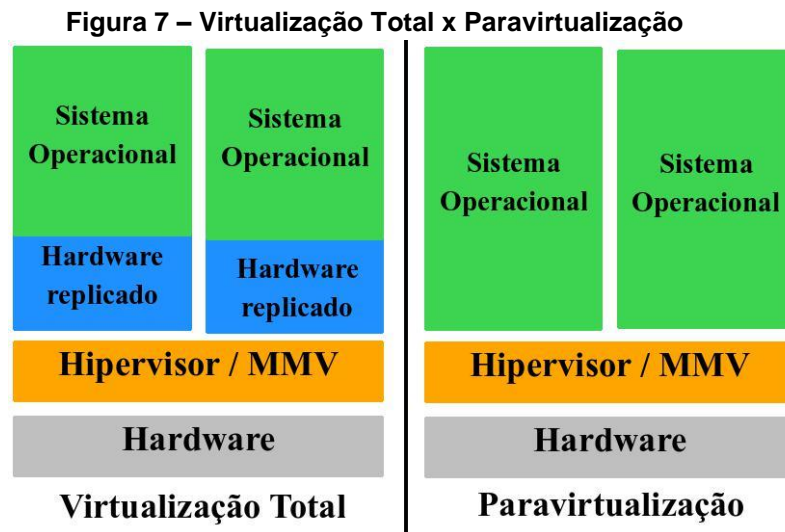
<https://www.devmedia.com.br/hypervisor-seguranca-em-ambientes-virtualizados/30993>

2.4.2 Tipos de Virtualização

Os tipos de virtualização de servidores, inicialmente definidos para a arquitetura x86 eram classificados em: (1) virtualização total; (2) paravirtualização/assistida pelo SO; (3) virtualização assistida pelo hardware.

A virtualização total é a abstração completa do hardware físico combinada com técnicas de translação binária (para lidar com os privilégios dos processos) e de execução direta. A camada de software é a responsável por controlar todo o hardware subjacente e por isso há uma perda de desempenho. O sistema operacional nativo e os convidados “têm a impressão” de estarem executando sobre o hardware original e não são necessárias modificações nem no SO nem nas VMs para suas execuções.

Na paravirtualização, os recursos de hardware são abstraídos mas não é uma abstração idêntica ao hardware físico. Os dispositivos de hardware são acessados por drivers do próprio hypervisor, mas os sistemas operacionais convidados devem ser adaptados para utilizarem os hypervisors sempre quando forem executar instruções sensíveis. As instruções de usuário (privilégio menor) podem ser executadas diretamente no processador. É uma forma de se resolver problemas associados à desempenho de execução utilizando-se da virtualização.



<https://www.portalgsti.com.br/2016/11/virtualizacao-completa-e-paravirtualizacao.html>

A virtualização assistida por hardware foi justamente o citado movimento feito pela Intel e AMD para alterarem os modos de operação do processador viabilizando a execução do *hypervisor* com total prioridade sobre o sistema operacional.

2.4.3 Conceito de *workload* e *throughput*

Os conceitos de *workload* e *throughput* são importantes no contexto da virtualização. *Wordload* ou carga de trabalho se refere ao conjunto de dados a serem processados a partir das requisições dos usuários. É variável pois em diferentes momentos do dia os usuários realizam mais ou menos requisições e é dependente da aplicação (o tipo de transações, consultas que são executadas).

O conceito de *throughput* está relacionado à capacidade de processamento: velocidade de leitura/escrita, da CPU, funcionalidades do processador, cache, quantidade de núcleos, etc.

2.4.4 Fornecedores conforme *Gartner*

A *Gartner* é uma empresa renomada de consultoria relacionada à pesquisa e benchmarking de soluções tecnológicas. A ferramenta de visualização de informação denominada quadrante-mágico (**FIGURA 6**) tem como objetivo fornecer análise qualitativa de mercado indicando principais participantes (fornecedores).

Figura 8 – Quadrante-mágico do Gartner para soluções de virtualização de servidores x86



GARTNER (2016)

2.4.5 Vantagens da Virtualização de Servidores

No período em que a infraestrutura de Data Center era orientada para o modelo cliente-servidor (Data Center 2.0 - **Figura 3**), o mais comum era que uma aplicação executasse em um equipamento servidor dedicado (1:1). Esse equipamento era geralmente superdimensionado em sua aquisição e era conseqüentemente subutilizado em sua capacidade de processamento e memória, por ter apenas uma aplicação consumindo seus recursos. Com o uso da virtualização, várias aplicações podem ser executadas em um único equipamento ou cluster de equipamentos, **aumentando o uso potencial dos equipamentos e conseqüentemente a eficiência do data center, reduzindo o número de servidores físicos necessários.** SANTANA (2013), aponta o uso de **15%** da capacidade dos servidores físicos sem virtualização. Com a virtualização há um aumento expressivo da utilização dos equipamentos que vai variar conforme a utilização do equipamento em cluster (em grupo) ou *standalone* (apartado). O planejamento da capacidade de utilização de um servidor deve considerar, em caso de grupo de servidores (cluster), que quando um dos nós falhar, os outros devem ser capazes de receber a carga de trabalho de forma dividida, para

evitar sobrecarga em um dos servidores. Em um exemplo de um cluster de dois servidores (equipamentos com características iguais denominados nós), o ideal seja que funcionem a **50%** da sua capacidade, pois em caso de falha de um dos nós o outro consiga assumir os 100% da carga de trabalho existente. Servidores standalone (sem estarem em grupo) virtualizados podem ter capacidade utilizada (memória e processamento total) média de 80% ou mais considerando o planejamento das aplicações que serão executadas no equipamento.

O já citado Instituto Brasileiro de Museus (IBRAM) possui quinze servidores físicos do mesmo modelo (**DELL R900**) obtidos através de doação no período de 2011 a 2014. Esses equipamentos apresentam, cada um, 128 GB de memória RAM e 34,08 Ghz de processamento (4 processadores * 4 núcleos * 2,13 GHz). As fontes redundantes dos equipamentos possuem uma potência nominal de 1570W. Até 2018, os quinze servidores eram utilizados ou de forma dedicada para poucas aplicações (sem virtualização) ou com o virtualizador Ovirt, que com o tipo de licença utilizada compatível com os equipamentos em questão não permitia a implementação correta de cluster de servidores. Após levantamento do uso dos recursos computacionais, verificou-se que a memória total utilizada pelas aplicações institucionais era de 768,95 GB RAM, com processamento máximo de 90 GHz. O órgão possui licença Windows Server 2012 que permite o uso do virtualizador Hyper-V. Um novo projeto de virtualização já implementado considerou **dez** dos quinze equipamentos configurados em cluster (1,28 TB RAM e 340 GHz de processamentos disponíveis considerando crescimento) utilizando o virtualizador *Hyper-V*, liberando cinco equipamentos para serem disponibilizados para unidades museológicas sob responsabilidade do IBRAM mas que são mais representativas (tamanho, quantidade de visitas, quantidade de acervo). A correta implementação de virtualização no órgão **permitiu a redução do consumo de energia (passaram a ser utilizados menos servidores para hospedar o mesmo conjunto de aplicações)**, com maior disponibilidade dos serviços pois quando configurados em cluster, em um *failover* (falha) de um nó, a máquina virtual é recriada em outro, havendo a continuidade dos serviços. Assim, o projeto conseguiu reduzir na sede da Instituição o consumo de energia em 33% (1 terço), assim como **permitiu liberar espaço de 20 Us** (cada equipamento servidor possui 4 Us) - metade do espaço de um rack, **reduzindo o espaço físico necessário para que os equipamentos hospedem as aplicações institucionais.**

No que tange à facilidade de alocação de recursos existentes no equipamento, outro benefício do uso da virtualização, é importante para isso a análise do consumo histórico das aplicações para dimensionamento apropriado da infraestrutura virtual a ser criada. A tecnologia de virtualização permite ainda assim **escalabilidade e flexibilidade na alocação** ao possibilitar o aumento dos quantitativos virtuais previamente definidos (memória, processamento, armazenamento, tráfego), e conseqüentemente

a resposta adequada à carga de trabalho e suporte a maior à acessos (usuários). A alocação é feita dinamicamente para aumento dos recursos. Na diminuição de recursos alocados é geralmente necessária a criação de uma nova máquina virtual com novas e menores dimensões.

A máquina virtual (VM) ao ser criada é **isolada** das outras máquinas virtuais uma vez que seus processos não interferem no contexto das outras VMs. Além da possibilidade mais simples e já citada de serem executados múltiplos sistemas operacionais em máquinas virtuais distintas, pode-se também executar aplicações não-confiáveis em VMs buscando-se análise de comportamento e isso pode ocorrer no mesmo ambiente de produção de aplicações estáveis. É possível também executar softwares legados a partir de VMs que simulem hardware compatível.

Os softwares gerenciadores de ambiente virtualizado centralizam várias funções de administração do ambiente (dos estados dos dispositivos físicos), dos servidores físicos, dos clusters configurados, e várias informações sobre cada VM. **Provêm interface única de gerenciamento e geralmente simplicidade de uso**. As VMs são arquivos ou volumes lógicos e por terem essa natureza são **facilmente copiadas (para realização de backup)**, assim como são **fáceis de serem migradas** entre hosts. Copiar uma VM e adaptá-la para particularidades também facilita o processo de implantação de soluções.

2.4.6 Desvantagens da Virtualização de Servidores

Os recursos físicos computacionais baratearam muito ao longo dos anos, porém quando se fala em limitações para virtualização em servidores ainda a memória é o principal elemento, por ser ainda cara dentro do contexto da solução.

No IBRAM está sendo conduzido processo de aquisição de equipamentos e dentre eles, servidores de rede corporativos. O fator mais importante a ser verificado é a memória dos equipamentos a serem comprados tendo em vista a sua utilização na virtualização. Existe orientação do Ministério da Economia que compras de ativos de infraestrutura devam possuir suporte e garantia de cinco anos para se evitar maiores gastos com contratos de suporte técnico/extensão de garantia que tendem a ser mais caros. Dessa forma, o crescimento organizacional e conseqüentemente do ambiente virtualizado deve ser estimado para esse período, conforme **TABELA 1**.

Período	Memória RAM em GB
---------	-------------------

Período	Memória RAM em GB
1º Ano	768,95
2º Ano	961,18
3º Ano	1.201,48
4º Ano	1.501,85
5º Ano	1.877,31

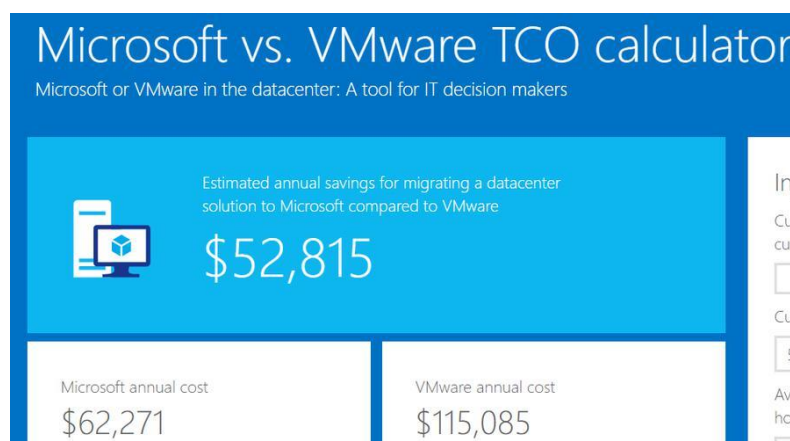
Fonte: Instituto Brasileiro de Museus

TABELA 1 - Memória total necessária como requisito para a contratação, considerando-se taxa anual de 25% de crescimento e o período de 60 meses (5 anos) a partir da memória atualmente utilizada pelas aplicações (Tabela extraída do estudo técnico da contratação dos equipamentos servidores de rede) .

Além da previsão correta de uso de memória e processamento, existem os custos de licenciamento e treinamento, que podem ser entendidos como desvantagens da utilização da virtualização. Em um breve exercício para estimar os custos de licenciamento e treinamento considerando os novos equipamentos a serem adquiridos no IBRAM, foram comparados os custos entre solução Microsoft (Hyper-V) e solução VMware. A Microsoft recentemente alterou o modelo de licenciamento que era por pro-cessador para núcleos (cores) tendo em vista a tendência de venda de equipamentos com menos processadores e mais núcleos.

Considerando que: (1) os cinco servidores de rede a serem adquiridos possuem cada: 1 processador com 12 núcleos - 2,1 GHz e 384 GB RAM; (2) atualmente existem no IBRAM 100 Máquinas Virtuais, cada uma com 2 VCPUs e 8 GB RAM de memória; e (3) seriam treinadas pessoas na solução de virtualização escolhida.

Figura 9 – Cálculo do licenciamento Microsoft anual, com base na infraestrutura citada. A ferramenta online disponível no site <http://datacenter-tco-tool.azurewebsites.net/> ainda compara com o preço anual da solução VMware.



Para verificação do preço da solução VMware de forma mais imparcial foram

utilizados os preços disponíveis no site da VMware (https://www.vmware.com/br/reusable_content/vsphere_pricing.html#): (1) 5 licenças vsphere Enterprise plus - U\$ 3.595,00 o suporte técnico (U\$899,00 por ano); (2) vCenter Server Standard - U\$ 6.175 e o respectivo suporte técnico (U\$ 1.544,00), totalizando U\$ 30.189,00 no primeiro ano. Existe uma disparidade entre o valor obtido pela ferramenta microsoft dos preços dos produtos VMware e o valor calculado com base em tabela online. Não é transparente os elementos utilizados no cálculo da ferramenta Microsoft para obtenção dos resultados. Esses valores ainda são valores de mercado. Nas compras feitas pelo governo os valores das licenças costumam ser maiores.

Calculando em valor médio na moeda nacional (considerando o preço U\$ 62.271,00 anual indicado pela ferramenta da Microsoft e o \$30.189,00 obtido de forma manual para o licenciamento VMware), existe **o gasto anual estimado de R\$ 184.920,00 com licenciamento dos produtos de virtualização** para os equipamentos citados no exemplo, um valor expressivo considerando as ferramentas líderes de mercado, e o dólar que está com cotação alta (U\$ 1,00 = R\$ 4,00 Reais).

Ainda neste exercício foram verificados os custos de treinamento nas respectivas soluções. O curso de “Configuração: VMware vSphere: instalação, configuração e gerenciamento (v 6.7) está R\$ 18.535,83 (Treinamentos VMware, 2019) por pessoa. Foi quotado o curso de “Instalação, armazenamento e computação com o Windows Server 2016” através da parceira credenciada Microsoft “KA Solution” associado com a utilização de Hyper-V, obtendo-se o preço de R\$ 1.255,00 por pessoa.

2.4.7 Considerações recentes

LEÃO CARNEIRO; MELO (2016), questiona em seu trabalho se um ambiente de TI que faz uso de virtualização possui vantagens em relação a um ambiente tradicional, considerando fatores como a disponibilidade de serviços, a melhor utilização de ativos de informática, gestão dos recursos, necessidades de treinamento e custos. O estudo comprova, através da virtualização de servidores do Data Center do Centro de Tecnologia da Informação e Comunicação (CTIC) da Universidade Federal do Amazonas (UFAM) que há aumento de disponibilidade dos serviços, que os recursos consolidados são melhor utilizados e gerenciados porém ressalta o problema crônico do fraco investimento em formação que afeta os departamentos de TI: pressupõe-se que haja pessoal capacitado para gerenciar os recursos virtualizados.

Ainda em LEÃO CARNEIRO; MELO (2016), cita-se parte do **Planejamento Estratégico** da Universidade de São Paulo (USP), de 2011:

No atual ambiente de TI o Data Center é um componente central de missão

crítica. É, portanto, indispensável que seja concebido para oferecer os serviços da Universidade de maneira flexível e dinâmica, acompanhando as tendências tecnológicas e de sustentabilidade hoje exigidas. Os projetos devem considerar a elaboração de infraestrutura física padronizada, escalável e segura, provendo níveis de redundância e virtualização.

O autor cita também que artigos foram publicados entre 2008-2010 desaconselhando o uso da virtualização para aplicações críticas como Banco de Dados; e **que com a evolução do hardware dos equipamentos e dos softwares de virtualização, essa preocupação deixou de fazer sentido.**

2.4.7.1 Virtualização e a utilização de Contêineres

As soluções de virtualização tradicionais operam na orquestração do uso dos recursos físicos dos equipamentos, divididos pelas máquinas virtuais criadas que possuem cada uma o seu respectivo kernel.

A tecnologia dos contêineres atua sobre um sistema operacional (um kernel definido), criando ambientes virtualizados com processos reduzidos e apenas as bibliotecas necessárias para a execução das aplicações. Dessa forma, reduz-se, a necessidade de VMs que constituem-se como ambientes completos para cada uma das aplicações. Aplicações que usam uma mesma infraestrutura podem coexistir por meio de containers que acessam o mesmo kernel.

A partir do lançamento oficial do Docker em 2013 as empresas vêm utilizando a tecnologia dos contêineres para promover o compartilhamento de kernel dos hosts das aplicações que passam a estar encapsuladas com um conjunto mínimo de processos de sistema. Google vem utilizando a tecnologia desde 2008. É importante dizer que a virtualização convencional (de equipamentos) e o uso de contêineres não são tecnologias concorrentes, podendo ser utilizadas de forma combinada para aumentar a resiliência e a disponibilidade do ambiente virtualizado, permitindo maior agilidade na criação de soluções de TIC.

Figura 10 – Virtualização tradicional com o uso do VMware vSphere versus o uso de Kubernetes como orquestrador de contêineres.



Fonte: <https://blogs.vmware.com/cloudnative/2017/10/25/kubernetes-introduction-vmware-users/>

Na **Figura 10** acima, exemplifica-se a utilização da virtualização tradicional com o VMware vSphere instalado em um servidor físico versus um nó da estrutura do contêiner que pode ser tanto uma máquina física como uma máquina virtual. Na orquestração de contêineres usando o Kubernetes, são denominados pods a estrutura mínima e executável contendo um conjunto coeso de funcionalidades.

O ambiente de contêineres conta com um conjunto grande de imagens pré-definidas para instanciação e isso agiliza o processo de implantação de novas aplicações. A forma de se trabalhar com a tecnologia, orientada a microserviços, permite definir contêineres com escopo de trabalho específico ao invés de uma estrutura grande e monolítica de software. Assim pode-se monitorar o uso de cada um dos componentes utilizados por uma aplicação para se entender os gargalos, otimizar desempenho adhoc e geral da aplicação. A tecnologia permite redução de custos de mão de obra pela facilidade na execução de atividades de TIC comuns e também na administração geral do Data Center de uma empresa, que passa a utilizar uma infraestrutura virtualizada mais enxuta.

2.5 Virtualização de Armazenamento

O uso da virtualização em armazenamento deve considerar vários elementos tais como tipos de licenciamento/appliances, a rede de armazenamento, os tipos de equipamento storages (que também devem ser homologados na solução), interfaces utilizadas e o conjunto de funcionalidades necessárias ao negócio organizacional, além de depender da possibilidade da consolidação do armazenamento. Assim, pela complexidade maior, serão discutidos os principais elementos do sistema de armazenamento de um Data Center.

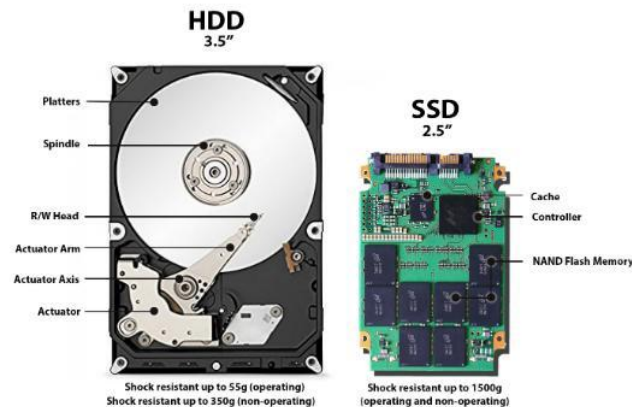
2.5.1 O Sistema de Armazenamento

O sistema de armazenamento de um Data Center possui três principais componentes: **(1) o equipamento de armazenamento (storage)**, componente principal da estrutura. Pode ser entendido como um servidor de discos e é responsável pelos requisitos de I/O (leitura/escrita) em um sistema computacional; **(2) os já citados servidores de rede**, responsáveis por processar as requisições das aplicações. Eles acessam os dados no equipamento storage **por blocos** (forma primária) **ou por arquivos** (abstração do modelo em blocos); **(3) conectividade**: a interconexão de rede necessária para que servidores e storage possam funcionar conjuntamente. Envolve cabeamento, equipamentos de rede, interfaces e protocolos.

A arquitetura do equipamento de armazenamento (storage) possui os seguintes principais elementos: **(1) discos**; **(2) controladora de discos**; **(3) cache**; **(4) portas ou interfaces**.

Atualmente existem discos mecânicos (hard-disk drives – HDD) e os discos de estado-sólido (SSD), baseados em memória não volátil NAND (memória do tipo flash capaz de reter informações em caso de perda de energia, ao contrário da Memória RAM). Os discos de estado sólido têm performances até 8x superiores às dos discos mecânicos. A discussão dos tipos existentes de discos e sua evolução se faz importante porque os processadores estão cada vez mais poderosos e isso está provocando um gap de desempenho entre os servidores de rede e os sistemas de armazenamento já que as operações de leitura e escrita nesses últimos equipamentos são mais lentas, provocando a ociosidade dos processadores. Também é importante citar que os discos mecânicos estão caindo em desuso com o barateamento das soluções SSD. Quanto ao uso, os discos SSD nos *storages* além de armazenamento mais performático, também vêm sendo utilizados como swap (área secundária de memória) para grandes cargas de trabalho (workloads). Substituiu-se assim a memória RAM de maior custos pelos discos SSD de menor custo.

Figura 11 – Discos Mecânicos vs SSD



Fonte: <https://www.backblaze.com/blog/ssd-vs-hdd-future-of-storage/>

A controladora é essencialmente um servidor (um processador junto com um controlador) responsável por realizar várias funções no equipamento tais como o gerenciamento de volume dos discos, o RAID (operações para maior confiabilidade e desempenho dos dados), controle do fluxo de dados, e mais atualmente funções como snapshots (estados instantâneos dos dados das aplicações), clones (cópias dos dados), thin provisioning (alocação dinâmica de recursos), auto-tiering (movimentação dos dados mais acessados para os discos de maior velocidade) e replication (replicação de dados em uma Storage Area Network). O cache da controladora desempenha papel importante na performance do controle de fluxo de dados e das operações.

A conexão dos storages com os servidores é feita através das tecnologias de Fibre Channel e iSCSI e essas permitem que os servidores de uma rede façam uso compartilhado dos recursos do equipamento, otimizando o uso de discos, isso dependendo do tipo de Rede de Armazenamento.

2.5.2 Os tipos de Storages

Os storages também podem ser divididos em três categorias: **(1) JBOD**: just a bunch of disks: subsistema de discos não possui controladora e o equipamento não possui funcionalidades como RAID (Redundant Array of Independent Disks); **(2) do tipo RAID**: storage com maior capacidade de proteção dos dados que é feita através de técnicas de replicação de dados. Aumenta-se o throughput do equipamento também ao se utilizar várias unidades de discos (que contém os dados) ao invés de uma só maior; **(3) storage inteligente**: os subsistemas de discos se utilizam de cache para acelerar o processo de leitura e escrita. A cache existe tanto nas controladoras como nos discos. Possuem além de RAID outras funcionalidades: capacidade de fazer snapshots, espelhamento remoto, lun masking (acesso autorizado a dados pelo servidor através de uma porta específica), a utilização de discos de substituição (hot spare) quando há

falha em algum disco: os dados são recriados no novo disco inserido a partir dos dados dos outros discos; e a desduplicação: capacidade de se economizar armazenamento de réplicas de dados utilizando-se apontamentos.

2.5.3 Tipos de Rede de Armazenamento

O aumento da quantidade de dados armazenados pelas aplicações provocou a mudança da localização desses dados que antigamente ficavam no próprio equipamento servidor da aplicação e passaram a estar nos equipamentos próprios de armazenamento. As formas de conexão dos servidores com os equipamentos storages também evoluíram: da conexão direta entre esses equipamentos (DAS – Direct Attached Storage) para a utilização de redes de protocolos e interfaces específicas (SAN e NAS).

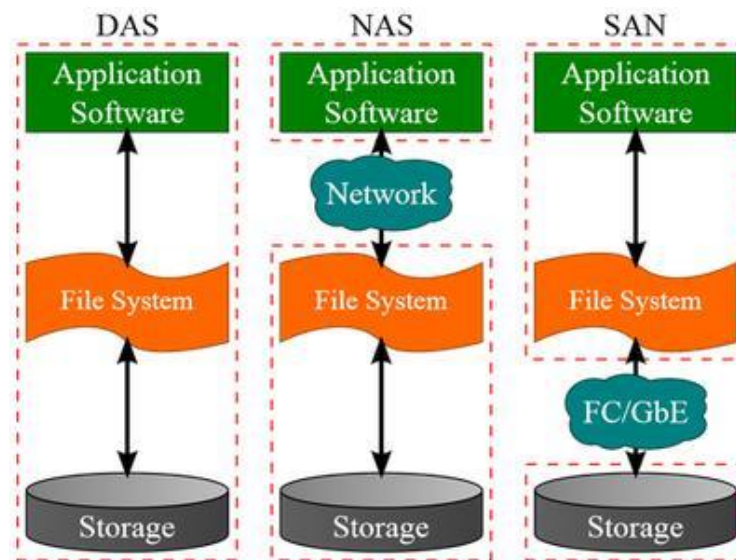
As redes SAN (Storage Area Network) conectam os servidores aos storages no nível de bloco (dados da aplicação). Se baseiam no padrão Fibre Channel (FC) ou Gigabit Ethernet (IP). O padrão Fibre Channel está consolidado no mercado, porém a implementação de uma SAN FC tende a ser mais cara. O padrão iSCSI sobre TCP/IP (Ethernet) simplifica o armazenamento e reduz o Custo Total de Propriedade (TCO) quando se compara as redes do tipo SAN FC e SAN IP. As redes SAN IP também permitem maior flexibilidade pois o padrão iSCSI pode ser incorporado em redes NAS e SAN FC existentes com a utilização de appliance e a existência das interfaces iSCSI (mais comum atualmente) nos dispositivos.

Redes do tipo NAS (Network Attached Storage) são utilizadas para o compartilhamento de arquivos e apresentam maior disponibilidade e escalabilidade do que as DAS. Utilizam o TCP/IP para transferência de dados e NFS/CIFS para serviço de arquivo remoto. Vários servidores e usuários podem acessar a mesma estrutura NAS que também trata questões envolvendo a permissão de acesso.

Em ambientes corporativos atuais é comum de se encontrar redes SAN e NAS para atender a propósitos distintos.

Hoje fala-se em convergência de redes e do sistema de armazenamento através do uso do padrão FCoE (Fibre Channel over Ethernet) capaz de integrar formas já consolidadas de conexão utilizando frames FC através de rede Ethernet. FCoE é alternativa ao iSCSI visando a integração do ambiente. O adaptador de rede convergente (CNA) permite a unificação das redes a partir das tecnologias Ethernet e FCoE. A convergência permite o barateamento da estrutura de rede do sistema de armazenamento de um Data Center: cabeamento, energia, resfriamento, gerenciamento e facilita a virtualização do ambiente.

Figura 12 – DAS, NAS e SAN - Tipos de Redes de Armazenamento



Fonte: <https://www.portalgsti.com.br/2017/10/snua-brasil-career-talk-1910.html>

2.5.4 Virtualização

A virtualização do armazenamento é a abstração lógica do sistema físico dos storages e passa a ser considerada (fazer sentido) em ambientes com equipamentos heterogêneos (compatíveis e homologados na solução de virtualização). A virtualização pode ocorrer nos discos dos servidores de rede (espaço de armazenamento geralmente subutilizado), no nível de bloco para rede SAN e arquivo para redes NAS. Inclui também, como componente, discos, unidades de fitas e sistemas de arquivos.

O administrador passa a ter um ponto central em que vê todo o espaço de armazenamento corporativo de forma unificada, facilitando o seu gerenciamento. Geralmente utiliza-se da virtualização de redes (caminhos separados para controle e dados) mas também pode ocorrer com implementações tradicionais a partir do uso de appliances.

2.5.5 Storage Definido por Software (SDS)

O conceito de SDS está relacionado com a automatização do armazenamento (através de regras da aplicação nos servidores - com provisionamento dinâmico) e em termos de arquitetura se baseia em hardware de tecnologia aberta capaz de ser administrado por diferentes tipos de hypervisors e ferramentas de orquestração baseadas em API (ex.: openstack).

2.6 Virtualização de Redes

2.6.1 Conceito

A virtualização de redes surgiu para superar os problemas das amarras das tecnologias proprietárias existentes nos equipamentos que habitualmente fazem parte de uma estrutura de redes. O hardware proprietário limita o estabelecimento/crescimento das redes por questões de custo e dificulta em termos de interoperabilidade a criação dos possíveis melhores arranjos topológicos visando a otimização da estrutura do Data Center.

2.6.2 Paradigmas

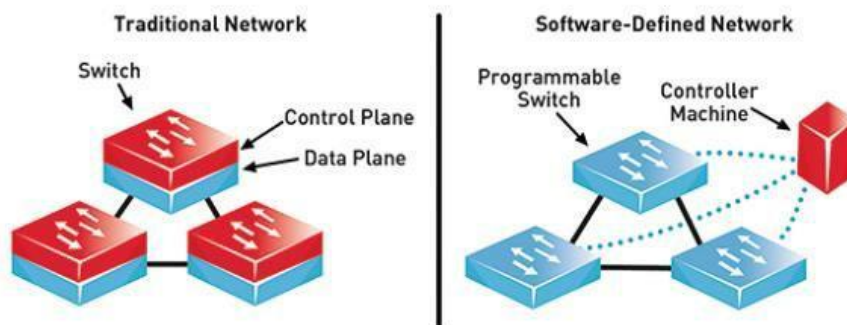
2.6.2.1 SDN

A Rede Definida por Software (SDN) se constitui um dos paradigmas de virtualização de redes que tem como característica-chave a separação da rede virtual da rede física por meio de uma camada de software que “esconde” a topologia das aplicações. Pode ser entendida também como uma plataforma independente de fabricante (dos switches, roteadores proprietários) e promove a separação útil entre o **controle das ações dos equipamentos e dos dados**.

O termo SDN surgiu quando do desenvolvimento do protocolo OpenFlow a partir de 2007, protocolo **que viabiliza a comunicação do plano de controle com o plano de dados**. A grande contribuição da SDN foi possibilitar a rápida configuração da rede conforme a demanda de negócios e serviços e possibilitar canais virtuais dentro da mesma rede de produção para testes de recursos de inovação sem impacto nos serviços operantes dentro dessa mesma infraestrutura. Assim, a SDN introduziu “programabilidade” à rede e ponto único de gerenciamento. Equipamentos de comunicação (geralmente virtualizados) encaminham regras de acordo com políticas que se encontram em servidores. Desta forma podem ser inseridas novas regras de controle.

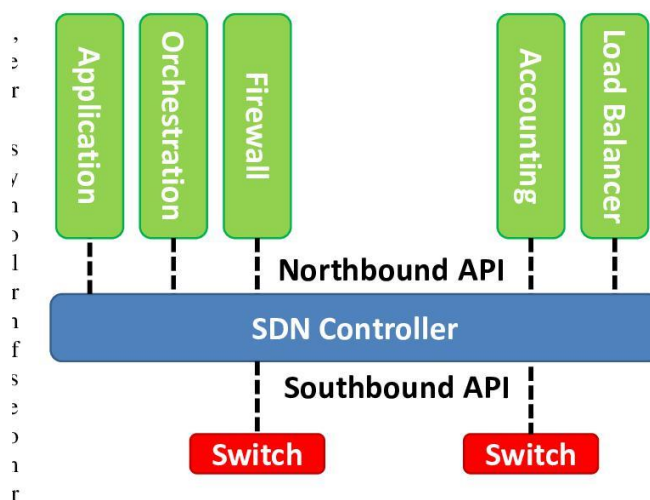
A SDN, por introduzir uma camada de software na rede é mais um ponto de brecha de segurança e por isso, deve existir a correta proteção no acesso ao Controlador SDN. Além disso, ela apresenta um processamento oneroso: cada comutador (elemento virtual encaminhador) se comunica constantemente com o controlador de rede o que pode gerar overhead (aumento do processamento de tráfego).

Figura 13 – Rede Tradicional vs SDN



Fonte: <https://penseemti.com.br/artigos/o-que-e-sdn-e-como-essa-arquitetura-de-rede-funciona/>

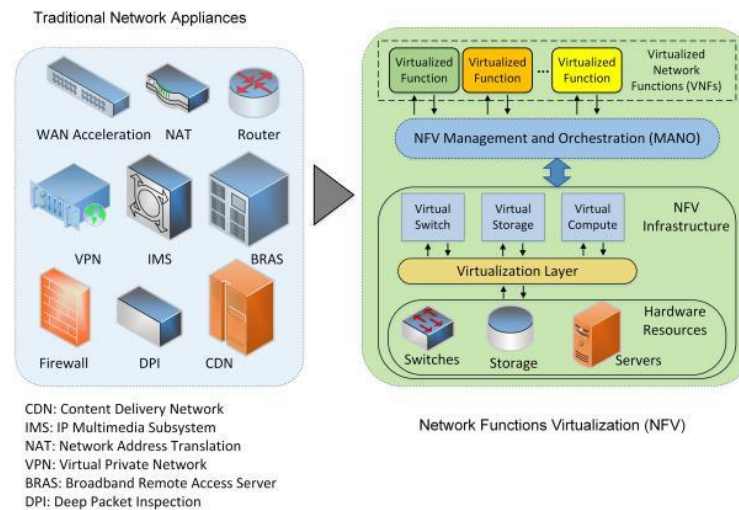
Figura 14 – Aplicações, Camada de Controle (SDN Controller), Camada de Dados (switch). A comunicação entre SDN Controller e os vswitches ocorre por meio do protocolo Openflow (forma mais comum).



Fonte: https://www.researchgate.net/figure/The-Southbound-and-Northbound-SDN-interfaces_fig2_27636934 2

2.6.2.2 NFVs (Funções Virtualizadas de Redes)

As funções de rede virtualizadas (NFV) se constituem outro paradigma de virtualização de redes, relacionadas à virtualização pontual de elementos da rede, utilizando-se de hardware genéricos (servidores) emulando em VMs com apoio de hypervisor funções como: nat, proxy, firewall, dns, dhcp, encaminhamento, roteamento, balanceamento de carga. As NFVs viabilizam a criação/crescimento de redes mais baratas, mas exigem um esforço de programação para manter o mesmo desempenho dos dispositivos físicos.

Figura 15 – *emph*Appliances de rede vs NFVs

Fonte: <https://www.sciencedirect.com/science/article/pii/S0167739X16302321>

2.6.3 Considerações recentes

MARCHESAN (2018), em seu trabalho de dissertação, compara os paradigmas independentes, porém complementares de virtualização de redes: Redes Definidas por Software (SDN - Software Defined Networking) e Virtualização das Funções de Rede (NFV - Network Function Virtualization) estabelecendo cenários em que um ou outro ou ambos sejam melhor utilizados. A dissertação ainda enfatiza a importância que os middleboxes - equipamentos appliances (hardwares dedicados) que realizam conjunto específico de funções de rede - desempenharam na consolidação das redes, porém por apresentarem custos altos de aquisição e manutenção e por possuírem tecnologias proprietárias dificultam atualmente a flexibilização das redes corporativas para atender às novas demandas de usuários. Em relação às SDNs, elas apresentam um processamento oneroso pois os elementos encaminhadores (virtualizados) - camada de dados - de fluxo da rede se comunicam intensamente com o controlador da rede (camada de controle) e precisam ser repensadas para certos tipos de aplicações. As NFVs surgem como alternativas para abstrair a camada física da camada de controle porém para funções específicas, possibilitando novos arranjos de rede com o uso de equipamentos genéricos utilizando hypervisors (executando em servidores virtualizados), contribuindo para a redução de custos de capital (CAPEX) e de operações (OPEX). Nesse mesmo trabalho foram testadas as funções de firewall, switching, roteamento e dhcp server de forma quantitativa (através de métricas de desempenho) e qualitativa (segurança, programabilidade, resiliência/disponibilidade) em ambos os paradigmas. Verificou-se que: NFVs são mais adequadas para funcionalidades de proteção de rede e SDN para funções de interoperabilidade e ambas funcionam eficazmente para otimização, monitoramento e controle das redes.

3 Materiais e Métodos

O tópico “Virtualização” surgiu, conforme dito na introdução deste documento a partir da necessidade de se facilitar a compreensão dos conceitos envolvidos na otimização de recursos de TIC, no ambiente de infraestrutura, na compra de equipamentos e na contratação de serviços, no entendimento de questões relacionadas ao desenvolvimento de sistemas. O componente “Virtualização” é meio para várias tecnologias do futuro tais como IoT; não menos importante, é consolidar as informações em um documento e atualizá-lo com artigos recentes. **Este é um trabalho de revisão bibliográfica.**

O trabalho foi pensado de forma a se estabelecer uma linha do tempo (cronologia) dos eventos relacionados as tecnologias de virtualização nos diferentes contextos com base em duas literaturas consolidadas: (1) um material nacional do Manoel Veras Sousa de 2016; (2) um material internacional do Gustavo Santana de 2013. Essas duas literaturas contribuíram com a elucidação dos conceitos desta Monografia.

O segundo passo foi procurar em site especializado do Google de artigos acadêmicos, artigos relacionados aos diferentes contextos de virtualização no período de 2016 a 2019. Foram obtidos artigos dos tipos de segmentos: virtualização de servidores, de armazenamento e de redes. Foram transcritos e explanados pontos importantes abordados pelos artigos.

O escopo do trabalho foi reduzido para Virtualização de Infraestrutura considerando apenas os contextos de servidores, armazenamento e redes em detrimento da virtualização de desktops, aplicações e nuvem por dois motivos: (1) prazo para elaboração da monografia que foi realizada em 3 (três) meses; (2) priorização dos contextos mais importantes considerando a realidade da Administração Pública Federal.

4 Resultados

A virtualização de servidores está sendo amplamente utilizada na Administração Pública. Discute-se nos órgãos, atualmente, a adoção da virtualização de armazenamento - quando há a existência de ambiente heterogêneo de equipamentos - e a virtualização de redes (SDN). As funções virtualizadas de rede já vêm sendo utilizadas: firewall, proxy, etc.

Existe uma preocupação na utilização dos equipamentos legados na composição do projeto de virtualização institucional, mas esse pensamento deve ser reconsiderado. Os componentes físicos presentes na infraestrutura do ambiente a ser virtualizado devem ser compatíveis com as soluções recentes de virtualização, devem possuir garantia e suporte técnico e/ou serem em número o suficiente para em caso de falha de um deles a estrutura permaneça operacional. Há a orientação da divisão do ambiente em legado (quando os equipamentos não podem ser utilizados na estrutura de virtualização) e no ambiente virtualizado. Não há como pensar em virtualização sem pensar em componentes de um Data Center (local ou recursos em nuvem).

Anteriormente ao projeto de virtualização, devem ser institucionalmente definidas as aplicações estratégicas/críticas e as táticas/operacionais para priorização do uso da infraestrutura mais resiliente para os tipos diferentes de ambiente (produção, desenvolvimento, testes).

No projeto da virtualização, projeto de média a alta complexidade, devem ser considerados: (1) equipamentos a serem interconectados, homologados nas soluções de virtualização e componentes/funcionalidades presentes dentro de cada equipamento; (2) forma de licenciamento e valor das soluções; (3) interfaces e protocolos a serem utilizados; (4) eliminação de possíveis pontos de gargalo de tráfego; (5) tipos de discos do storage e funcionalidades de armazenamento disponíveis; (6) capacitação dos responsáveis pelos serviços virtualizados; (7) devem ser calculados os diferentes custos totais de propriedade (TCOs) para os diversos cenários encontrados buscando atendimento da necessidade organizacional e economicidade; (8) aumentar o uso geral dos equipamentos do Data Center ; (9) planejamento das ordens de serviço dos serviços de sustentação de ambiente de infraestrutura terceirizados possibilitando a eliminação de procedimentos manuais que passarão a ser virtualizados e a incorporação de atividades relacionadas à gestão dos recursos virtualizados.

5 Conclusão

A virtualização é ferramenta valiosa na otimização dos recursos de um Data Center, mas sua correta utilização depende de um projeto bem estruturado e orientado estrategicamente para o negócio. Conhecer sobre todos os conceitos envolvidos viabiliza o diálogo interno e com fornecedores na busca por melhores soluções de TIC organizacionais.

A maturidade dos órgãos públicos federais ainda é baixa ou média o que ainda exige um esforço na virtualização dos recursos de TIC. O desenvolvimento da maturidade em virtualização permite a utilização de se adotar, por exemplo, nuvem pública para os serviços críticos, mitigados os riscos de segurança envolvidos.

O profissional de TIC para melhor projetar a topologia de rede e estrutura do Data Center deve conhecer de seus componentes físicos: tipos de cabeamento, elementos importantes em cada equipamento, como quantidade de processadores, processamento total possível de um servidor e a quantidade de memória RAM disponível; a capacidade de *throughput* das controladoras do *storage*, a existência de cache, os tipos de discos; os equipamentos existentes na rede, se são homologados nas soluções de virtualização. As interfaces de todos os dispositivos devem permitir interconexão e os dispositivos devem ser capazes de serem gerenciados pelas mais atuais plataformas abertas e funcionarem com protocolos como o *Openflow*.

O caminho para a virtualização de um Data Center envolve um trabalho de mapeamento dos recursos de TIC existentes, de segregação de ambientes (antigo/novo) com o descarte progressivo de equipamentos que não têm mais utilidade dentro da infraestrutura estabelecida. É importante a consciência do gestor de TIC que a maior parte das operações estratégicas e administrativas dos órgãos estão em aplicações e a estrutura física deve ser resiliente (contratos de suporte e manutenção e redundância de equipamentos). A virtualização nesse contexto é forma de se adquirir maior flexibilidade para crescimento, inovação, custos menores e agilidade na prestação do serviço público e na transformação digital pretendida pelo Governo.

O governo está começando a criar infraestrutura em nuvem híbrida (parte dos seus recursos de infraestrutura mais atual compatível com as tecnologias recentes) capaz de provisionar recursos localmente e de também migrá-los para infraestrutura (nuvem) externa em casos necessários. Softwares de gerência de virtualização (que possuem APIs com os diferentes *hypervisors*) como o *openstack*, e o *vCloud suite* estão sendo inicialmente adotados nesse contexto como ferramentas integradoras e orquestradoras dos diferentes tipos de virtualização explorados neste trabalho.

REFERÊNCIAS

- 1) MINISTÉRIO DO PLANEJAMENTO, DESENVOLVIMENTO E GESTÃO (MPDG). **Analista em Tecnologia da Informação**. Revisado em 03 de outubro de 2018. Disponível em: <http://www.planejamento.gov.br/acesso-a-informacao/recursos-humanos/cargos-e-carreiras-do-mp/analista-em-tecnologia-da-informacao>. Acesso em 10 de fevereiro de 2019.
- 2) MINISTÉRIO DO PLANEJAMENTO, DESENVOLVIMENTO E GESTÃO (MPDG). **Transição de Governo 2018-2019 - Informações Estratégicas**. Novembro de 2018. Disponível em: https://transicao.planejamento.gov.br/wp-content/uploads/2018/11/Informa%C3%A7%C3%B5es-Estrat%C3%A9gicas-Minist%C3%A9rio-do-Planejamento_vers%C3%A3o-publica%C3%A7%C3%A3o_completa.pdf. Acesso em: 04 de março de 2019.
- 3) BRASIL. PRESIDENCIA DA REPÚBLICA - CASA CIVIL. **Medida Provisória nº 726. 12 de maio de 2016**. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2016/Mpv/mpv726.htm. Acesso em: 04 de março de 2019.
- 4) SANTANA, Gustavo A. A. **Data Center virtualization fundamentals: understanding techniques and designs for highly efficient data centers with cisco nexus, ucs, mds, and beyond**. Cisco Press, 2013.
- 5) WIKIPEDIA CONTRIBUTORS. **Software-defined data center**. Revisado em 13 de setembro de 2018. Disponível em: https://en.wikipedia.org/w/index.php?title=Software-defined_data_center&oldid=859313753. Acesso em: 15 de janeiro de 2019.
- 6) SOUSA NETO, Manoel Veras. **Virtualização: tecnologia central do datacenter**. 2ª ed. BRASPORT, 2016.
- 7) BUENO, Henrique. **Virtualização: um pouco de história**. 29 de abril de 2009. Disponível em: <https://hbueno.wordpress.com/2009/04/29/virtualizacao-um-pouco-de-historia/>. Acesso em: 15 de janeiro de 2019.
- 8) DEVMEDIA. **Virtualização de servidores**. Disponível em: <https://www.devmedia.com.br/virtualizacao-de-servidores/30820>. Acesso em: 29 de janeiro de 2019.
- 9) SEO, Carlos Eduardo. **Virtualização: problemas e desafios**. Disponível em: <http://www.ic.unicamp.br/~ducatte/mo401/1s2009/T2/todos.pdf>

- 10) Nathuji, R. & Schwan, K. **VirtualPower**: coordinated Power management in virtualized enterprise systems. SOSP '07: Proceedings of twenty-first ACM SIGOPS symposium on operating systems principles. ACM, 2007, p. 265-278.
- 11) LEÃO CARNEIRO, João Bosco; MELO, Daniel Reis Armond de. *Análise da tecnologia de virtualização de servidores em um data center como fator para obtenção de ganhos de produtividade*. Revista **Espacios**. v. 37, nº 27. 2016. p. 20-32.
- 12) MARCHESAN, Gabriel. **Uma análise comparativa entre paradigmas de virtualização de redes**. Dissertação (Mestrado em Ciência da Computação). Universidade Federal de Santa Maria. Rio Grande do Sul. 2018. 122 p.
- 13) Treinamentos VMware. **VMware vSphere**: instalação, configuração e gerenciamento [V6.7] – Sob demanda. Disponível em:
https://store.vmware.com/store/e/vmwbr/pt_BR/DisplayProductDetailsPage/ThemeID.38316100/productID.5265860000. Acesso em: 21 de junho de 2019.
- 14) Sérgio Leandro. *Contêineres ou máquinas virtuais, qual é o melhor método para virtualização?* **Computerworld**. Disponível em:
<https://computerworld.com.br/2017/11/27/c-ontaineres-ou-maquinas-virtuais-qual-e-o-melhor-metodo-para-virtualizacao/>. Acesso em 21 de junho de 2019.