GERENCIAMENTO E PROTEÇÃO DE DADOS "ON PREMISE" E "CLOUD"

Autor: Rafael de Moura Campos

Resumo

A proteção de dados (backup) é parte vital de todo ecossistema tecnológico, os dados precisam ser guardados em pelo menos duas cópias feitas em locais distintos. A má gestão destes dados pode levar uma empresa ao fracasso. Com avanço das tecnologias de virtualização de datacenters, computação em nuvem e serviços de alocação de infraestrutura de datacenter, a tarefa de proteção de dados e gerenciamento de cópias de segurança se tornaram um desafio ainda maior. A demanda para utilização de serviços em computação em nuvem tende a aumentar significativamente nos próximos anos, tanto como serviço primário de infraestrutura como para contingência e plano de recuperação de desastres. Entre as empresas que oferecem software para proteção de dados, incluindo em computação em nuvem, existem duas que destacaram nos últimos anos e são consideradas líderes neste seguimento. A utilização de normas e pesquisas de mercado são essências para uma boa implementação e administração de cópias de segurança. Algumas empresas e pessoas costumam negligenciar este tema tão sensível e ao mesmo tempo essencial para a continuidade de um negócio.

Palavras-chave: Backup, Cloud, Proteção de Dados.

Data Management and Protection On-Premises and Cloud

Abstract

The data protection is vital for a whole technological ecosystem, the data needs to be backed up in two different copies in distinct places. The lack of management on those data can bring the whole company to fail. With the evolvement and improvements on the last years on, datacenter virtualization, cloud computing and colocation services, the data protection and backup management became a big challenge to most of the organizations. The cloud computing utilization trends are should a relevant increase in the next few years, especially in two main areas, infrastructure as a service a disaster recovery plan. Among all the companies that offer data protection and backup solution on the global market, there are two that became leaders of this competitive segment. The best practice and benchmark utilization are key factor to implement and manage backup environment. Some companies and people yet overlook this topic, which at the same time is considered such important process a part of business continuity.

1. INTRODUÇÃO

Em tempos atuais, onde o mundo está cada vez mais conectado e onde empresas precisam ser cada vez mais ágeis e eficientes, para que possam lograr êxitos em seus negócios, a proteção de dados (backup) é parte vital de todo ecossistema tecnológico. Desde um simples smartphone, onde seu usuário utiliza-o para transações bancárias, armazenamento de fotos, entre outras funções, até dados financeiros ou sigilosos de milhões de clientes de uma grande multinacional, os dados precisam ser guardados em pelo menos duas cópias feitas em locais distintos, para proteger os dados críticos da organização contra perda de dados que podem ser causadas por desastres ou softwares mal-intencionados. A má gestão destes dados e suas cópias pode levar uma empresa ao fracasso (MORAES, 2007).

Segundo Moraes (2007), há algum tempo o backup simplesmente significava cópia de segurança. Entretanto, no ambiente de Tecnologia da Informação, o backup e a proteção dos dados são utilizados para prover continuidade dos negócios. Porém com o advento da cloud computing e serviços de colocation[1] as práticas existentes de proteção de dados podem não ser tão eficazes, pois os dados podem estar em qualquer lugar. De acordo com Domingues, Denilson (2012) a virtualização de servidores de um datacenter pode facilitar o gerenciamento da infraestrutura, mas pode dificultar o gerenciamento de seus backups.

A norma ISO/IEC 17799 (ABNT, 2005) recomenda que cópias de segurança dos backups primários sejam feitas e armazenadas em locais remotos (off-site), com uma distância suficiente para que em caso de desastre no local primário, onde as cópias primárias serão perdidas, se possa evitar danos nas cópias secundárias. Atualmente ainda é muito comum encontrar empresas que utilizam fitas magnéticas para a cópia secundária, assim as mesmas são enviadas para um outro local, seja ele um outro prédio da empresa ou até mesmo salas cofres de empresas especializadas em proteção de dados.

As fitas magnéticas para proteção de dados consistem em dispositivos capazes de armazenar grandes quantidades de dados. Essas fitas magnéticas são de material plástico revestida de fina camada de material magnetizável, em que são gravados os dados. Para acessar uma determinada informação, é necessário ler todas as informações gravadas antes (BODÊ, 2009).

A cloud computing, é um novo modelo de computação que permite ao usuário final acessar uma grande quantidade de aplicações e serviços em qualquer lugar e independentemente da plataforma (SILVA, F. H. R., 2010) e com avanço das conexões de rede WAN (Acrônimo em inglês para Wide Area Network), hoje é possível substituir as fitas magnéticas por replicação de dados online para outro local, afim de protegê-los contra desastres e garantir a continuidade do negócio. Para um menor tráfego de rede na replicação desses dados, existem tecnologias de compressão e desduplicação capazes de reduzir significativamente o volume de dados armazenados. De acordo com DELL/EMC Brasil, desduplicação de dados busca redundância de sequências de bytes em grandes períodos de comparação. Sequências de dados (com mais de 8 KB) são comparadas com o histórico de outras sequências. A primeira versão de uma sequência armazenada exclusivamente é referida, em vez de ser armazenada novamente. Esse processo é

[1] Colocation: E a modalidade de alocação de um espaço físico dentro de um datacenter de fornecedores ou terceiros, que ofereça a infraestrutura para o funcionamento e operação dos ativos de TI da contratante.

inteiramente ocultado de usuários e aplicativos, de modo que o arquivo inteiro é legível depois de gravado.

Este artigo tem o propósito de alertar e orientar profissionais e gestores do setor de tecnologia da informação no assunto, pois ajudará a responder três perguntas que podem ser essenciais na hora de planejar e implementar uma solução de proteção de dados e copias de segurança, que são elas:

- Qual é a tendência tecnológica na área de backup e proteção de dados?

- Quais são as companhias que se destacam no mercado atual de backup e proteção de dados?

- E quais são as melhores práticas para se gerenciar dados descentralizados e suas cópias de segurança (backup)?

No próximo tópico será abordado, uma breve pesquisa no tema que abrange as tendências relacionadas ao assunto e como as empresas estão se preparando para estas novas demandas e desafios. Em seguida será dado uma breve abordagem de dois fabricantes que ganharam destaque no mercado mundial no tema de proteção de dados, por demonstrarem capacidade de entregar e executar com excelência a proposta de proteção de dados e trazer inovação ao integrar nativamente seus produtos e softwares de proteção de dados a plataformas de computação em nuvem. Por último, serão citadas as melhores práticas para implementar proteção de dados, onde os dados possam estar espalhados em diferentes localidades, incluindo em datacenters de fornecedores de serviços de computação em nuvem.

2. O DESAFIO DA PROTEÇÃO DE DADOS

Segundo o Gartner (2016), 30% das organizações de médio porte irão alavancar ainda mais o uso de computação em nuvem no modelo IaaS (Infrastructure as a Service) fazendo com o que a busca por soluções unificadas e de multi-plataformas de proteção e recuperação de dados aumente no mesmo ritmo. A necessidade de proteção de dados não está mais limitada em proteger os dados que se encontram dentro de datacenter privado das empresas (on premises data), mas também ao dado que se encontra em aplicações rodando em serviços como SaaS e IaaS (Software as a Service e Infrastructure as a Service) e com isso trazendo complexidade ao gerenciamento e dificultando a restauração do dado. Outra demanda prevista pelo Gartner (2016), está diretamente ligada ao crescimento do uso da computação em nuvem como contingência ou plano de recuperação de desastres, enviando seus dados através de replicação ativa diretamente para a nuvem. O número de empresas em busca dessa solução irá dobrar até o final do ano de 2018, comparado com o ano de 2016, onde 11% das corporações buscaram soluções nesta área.

O próprio Gartner faz uma comparação entre os principais competidores no mercado mundial, nos quais oferecem soluções de Proteção e Recuperação de Dados em Datacenter, onde estes competidores são ranqueados e divididos em quatro quadrantes diferentes; líderes (Leaders) que são aquelas empresas líderes do segmento e possuem capacidade de levar adiante as promessas de entrega para a tecnologia em questão; desafiantes (Challengers) que seguem logo atrás do líderes do seguimento, possuem capacidade de execução, mas apenas possuem uma parcela do mercado; visionárias (visioners) possuem um forte apelo na área de pesquisa e

desenvolvimento do tema abordado, porém as vezes não possuem o domínio da tecnologia ou não são capazes de executar e entregar o prometido; e por último as empresas denominadas concorrentes de nicho de mercado, que são as empresas que possuem domínio em um mercado específico dentro do mesmo tema, como por exemplo, proteção de dados para ambientes de TI virtualizados.

Figura 1 – Gartner Magic Quadrant for Data Center Backup and Recovery Software (Quadrante mágico da Gartner Group para Proteção e recuperação de dados em Datacenter).

Fonte Gartner Group, 2017.

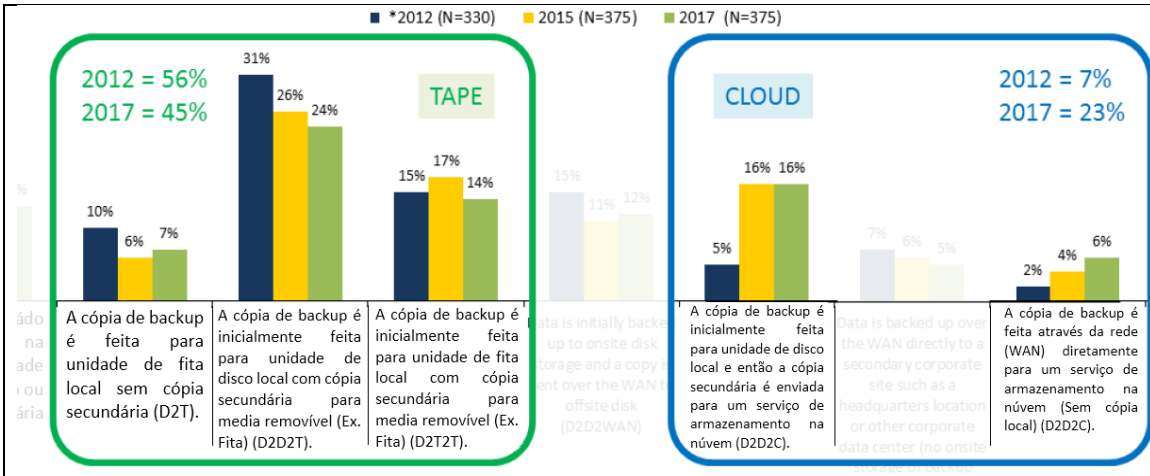
Conforme publicado por ESG Group (2016) os dois mais frequentes itens que foram reportados pelas organizações, como sendo os motivadores para buscarem a utilização de computação em nuvem para armazenamento de dados são: proteção de dados em sítios remotos e recuperação de dados em caso de desastres. Estes dois itens por sua vez acabam direcionando a organização para uma modernização da estratégia de gerenciamentos de dados, onde antigas premissas e orientações (melhores práticas) são substituídas por novas, como exemplo a utilização de fitas magnéticas, que anteriormente eram amplamente utilizadas para políticas de retenção de dados de longo prazo, estão sendo substituídas por replicações em tempo real para localidades remotas de provedores de serviços de computação em nuvem. Com a modernização da TI através de processos de negócio e governança baseada em práticas e normas de mercado, como por

exemplo ABNT NBR ISO/IEC 17799 (2005), e a busca por eficiência, tem motivado as empresas a buscarem novas tecnologias como, desduplicação, redução e otimização de cópias de segurança para sistemas não críticos vem sendo constatemente aplicadas como alternativas para retenção de dados e substituição de tecnologias obsoletas e pouco eficientes. De acordo com ESG Group (2016), nos últimos cinco anos (2012 – 2017) a utilização de arquitetura de backup baseada em fitas magnéticas caiu de 56% para 45% (11%) dentre as empresas participantes da pesquisa. Por outro lado, a procura por soluções de computação em nuvem cresceu de 7% para 23%, de acordo com as figuras 2 e 3.

Figura 2 – Principais motivadores para empresas procurarem ou optarem pelo uso de serviços infraestrutura de TI em nuvem.

Fonte Enterprise Strategy Group, 2016.

Figura 3 – Gerenciamento de proteção de dados 2012, 2015 e 2017.



Fonte Enterprise Strategy Group, 2016.

3. COMPARATIVO ENTRE OS LÍDERES DE MERCADO DE PROTEÇÃO DE DADOS

Conforme citado acima nesta sessão foi feito um comparativo entre dois fornecedores de ferramentas de proteção de dados que aparecem dentre os cinco fabricantes que se encontram no quadrante dos líderes feito pelo Gartner Group, conforme Figura 1. Estes dois fornecedores foram

escolhidos pois o autor teve acesso as ferramentas e ambientes de testes, onde foi possível obter conhecimento e dados suficientes para que a comparação fosse realizada.

3.1 Commvault

A Commvault é uma empresa norte-americana fabricante de software de proteção de dados que até a sua última versão 10 chamava-se Simpana Software, porém por estratégia de marketing o nome do software passou a ser o nome da própria empresa, a partir das versão 11, ou seja, Commvault®.

A Commvault surgiu em meados do anos 90, a partir da separação de duas gigantes do mercado de tecnologia na época, Lucent e At&T (Commvault 2016). Além de estar no topo da lista e liderar o quadrante mágico da Gartner Group, ela está na vanguarda da tecnologia de proteção e recuperação de dados e se destaca por ser uma solução multi-plataforma, ou seja, possui suporte a diversas plataformas de sistemas operacionais e banco de dados, o que torna o seu software de fácil implementação e altamente suportado e recomendado por diferentes fornecedores.

As principais características tecnológicas oferecidas pela Commvault são, compatibilidade com a múltiplas plataformas que vão desde diferentes fabricantes de banco de dados, de sistemas de virtualização até diferentes versões e fornecedores de sistemas operacionais. Sua arquitetura de desduplicação de dados através de políticas globais pode reduzir em até 90% a utilização de espaço nos dispositivos destinados a armazenamento dos backups (discos ou fitas magnéticas).

A Commvault possui uma integração já nativa na ferramenta para integração e proteção de dados na cloud e oferece compatibilidade com os principais fornecedores deste serviço.

Em contrapartida, segundo o Gartner Group (2017) a ferramenta Commvault pode ser relativamente complexa para se implementar e em seguida administrar, onde treinamentos oficiais são necessários e imprescindíveis para uma boa implementação e administração.

3.2 Veeam Backup & Replication

A empresa Veeam surgiu em meados do ano de 2006 onde lançou no mercado tecnológico seu primeiro produto, o Veeam Monitor que tinha apenas a funcionalidade de monitorar ambiente virtuais baseados na plataforma de virtualização do fabricante Vmware e somente em 2009 a Veeam lançou sua plataforma de backup, porém ainda era exclusivo para ambientes de virtualização da plataforma da Vmware. Somente em 2015 veio a estender a solução de backup e replicação para servidores físicos em duas plataformas de sistemas operacionais diferentes, distribuições Linux e Microsoft Windows.

O software Veeam Backup & Replication como é chamado, ganhou destaque no mercado de proteção de dados nos últimos anos por sua facilidade de implementação e gerenciamento, e por trazer confiabilidade no gerenciamento das cópias de segurança. O software ganhou popularidade no mundo tecnológico a medida que as empresas estavam cada vez mais investindo em datacenters virtuais, este produto trouxe a Veeam para o seleto grupo dos maiores fabricantes de softwares de backup e proteção de dados no mundo, e de acordo com Gartner Group (2017) ela se encontra na quarta posição atualmente em comparação de receitas.

A Veeam atualmente oferece uma compatibilidade com serviços de proteção de dados e replicação para cloud, sendo possível replicar o backup de seus sistemas críticos direto para um repositório na nuvem para que possa ser utilizado como estratégia de proteção contra desastres.

A solução oferecida pela Veeam é amplamente utilizada por empresas que possuem um parque computacional altamente virtualizado e que utilizam plataformas de virtualização compatíveis. Entretanto, para sistemas que ainda estão rodando em servidores físicos em diferentes áreas de um datacenter em plataformas legadas ou que não possuem compatibilidade com a ferramenta da Veeam precisarão ter seus dados protegidos de alguma outra maneira, trazendo assim a necessidade de uma segunda ferramenta de backup que precisará ser implementada e administrada pela equipe de gerenciamento de dados, trazendo assim um aumento nos custos de operação de TI e datacenter. Segundo o Gartner Group (2017) a Veeam ainda vai precisar de um tempo para comprovar a sua eficácia em sistemas físicos, para só aí então conquistar mais espaço neste competitivo mercado.

4. MELHORES PRÁTICAS DE IMPLEMENTAÇÃO E GESTÃO DE CÓPIAS DE SEGURANÇA E PROTEÇÃO DE DADOS.

As práticas que serão abordadas a seguir foram extraídas da norma ABNT NBR ISO/IEC 17799 (2005) e resumidas pelo autor. O mesmo resumiu aquelas consideradas mais relevantes na prática de proteção de dados e cópias de segurança.

- Os riscos relevantes a segurança da informação precisam ser identificados, quantificados e priorizados de acordo com a atividade da organização e seus objetivos.

- Classificação da informação – é conveniente que a informação seja classificada para indicar a necessidade, prioridades e o nível esperado de proteção quando do tratamento da informação. É recomendado que a informação seja classificada em termos de seu valor, requisitos legais, sensibilidade e criticidade para a organização.

- Papeis e Responsabilidades – Convém que papéis e responsabilidades pela segurança da informação de funcionários, fornecedores e terceiros sejam definidos e documentados de acordo com a política de segurança da informação da organização.

- Cópias de Segurança das informações relevantes - Convém que as cópias de segurança das informações e dos softwares sejam efetuadas e testadas regularmente conforme a política de geração de cópias de segurança definida.

- É imprescindível que rotinas de cópias de segurança sejam criadas para assegurar a integridade das cópias.
    - A frequência com que as cópias sejam feitas precisa atender os requisitos do negócio da organização e classificação da informação.
- Criar procedimentos operacionais padrão para que as equipes de operação de TI possam administrar e atuar nas rotinas de cópias de segurança em caso de falhas.
- Validar e testar as rotinas de recuperação das informações ou criar um calendário para que testes sejam executados regularmente.
- Criar procedimentos de recuperação da informação em caso de desastre e priorização das mesmas de acordo com a classificação de cada uma.

- Definir as políticas globais de retenção das cópias de segurança, atentando-se para os requisitos legais da natureza dos negócios da organização.

- Desenvolvimento e implementação de planos de continuidade relativos à segurança da informação – É importante que os planos sejam desenvolvidos e implementados para a manutenção ou recuperação das operações e para assegurar a disponibilidade da informação no nível requerido e na escala de tempo requerida, após a ocorrência de interrupções ou falhas dos processos críticos do negócio.

- Os tempos de recuperação das informações precisam ser definidos de acordo com a relevância de cada uma – "RTO" (Sigla em Inglês – Recovery Time Objective), que é referente ao tempo necessário para se recuperar a informação e disponibilizá-la ao negócio da organização. "RPO" (Sigla em inglês – Recovery Point Objective), que é referente ao tempo aceitável de perda da informação pelo negócio da organização.

A norma ABNT NBR ISO/IEC 17799 (2005) abrange diversas outras áreas referentes a segurança da informação, porém as citadas acima foram consideradas as mais relevantes para proteção de dados.

5. CONCLUSÃO

O mercado de gerenciamento e proteção de dados está em plena ascensão, tanto para as empresas quanto para pessoas que buscam mais agilidade no acesso aos dados críticos, e também para mantê-los protegidos de terceiros ou de grupos mal-intencionados. As empresas que já ofereciam sistemas de cópias de segurança tiveram que se adequar às novas demandas e tecnologias do mercado, como por exemplo computação em nuvem e desduplicação, pois hoje em dia a tecnologia da informação e consequentemente, os dados, são vistos como essenciais para o sucesso de uma empresa. Sendo assim, os dados precisam cada vez mais serem classificados e protegidos de acordo com sua importância para o negócio. A partir daí a equipe de TI e de administração de dados poderão desenhar a solução de proteção de dados e backup, para que seja uma solução onde, a eficiência pode ser facilmente demonstrada ao se comparar o custo total da mesma e subtraindo o valor dos dados que ela mesma está protegendo (Custo Benefício).

As duas tecnologias demonstradas no presente estudo são utilizadas por diversas empresas ao redor do mundo, essas tecnologias oferecem diversas funcionalidades que podem trazer grandes benefícios ao se tratar de proteção, classificação e gerenciamento dos dados, porém nenhuma dessas funcionalidades serão eficazes caso não haja controle e governança sobre a proteção de dados. Processos precisam ser previamente estabelecidos, políticas de retenção e expiração de dados devem ser implementadas e difundidas para clientes internos e externos. Acordos de níveis de serviço são essenciais, e também planos de recuperação contra desastres precisam ser criados e os seus tempos de recuperação devem ser acordados com os clientes e proprietários dos dados. Pessoas precisam ser altamente treinadas e conhecerem o ambiente, ter acesso aos planos de recuperação e classificação dos dados, para que em caso de necessidade possam agir autonomamente para tentar cumprir os prazos acordados, seguindo as a práticas citadas no tópico anterior.

Com relação ao uso de recursos de computação em nuvem, é importante ressaltar que é uma tendência global, e que a escolha pelo fornecedor dos serviços, e ferramenta que será usada para o gerenciamento das cópias de segurança e replicação, são essenciais para êxito da implementação. As duas empresas citadas neste artigo são capazes de integrar os recursos em nuvem e em datacenters locais (on-premise), porém é necessário sempre olhar para o custo benefício da solução comparando os valores das informações a serem protegidas.

Algumas empresas e pessoas costumam negligenciar este tema tão sensível e ao mesmo tempo essencial para a continuidade de um negócio. Este artigo teve como objetivo alertar e elucidar a importância de possuir ferramentas e processos robustos para proteção dos dados de uma empresa e até mesmo os nossos dados pessoais, pois em um tempo onde a tecnologia da informação faz parte do dia a dia da maioria das pessoas que vivem em uma sociedade moderna, os dados são as informações que criamos, trocamos, transmitimos e analisamos, sem eles toda a tecnologia ao redor não teria função alguma.

Este artigo teve também como objetivo trazer este tema para as mesas de discussões do mundo acadêmico, e não somente para estudantes e pesquisadores de tecnologia da informação, pois a tecnologia está em todas as áreas da sociedade e abrange quase todas as áreas acadêmicas. Ao se formar um profissional, o mesmo deverá estar atento ao tema de proteção de dados e a partir deste ponto, deverá tomar medidas preventivas ao seu alcance para aumentar a segurança e integridade dos dados que ele gera, acessa ou transmite.

6. REFERÊNCIAS

BODÊ, E. C. - PRESERVAÇÃO DE DOCUMENTOS DIGITAIS: RESUMO DA TEORIA. Brasília, 2009. Disponível em - <http://ebod.com.br/docsfrom/teoriacurso2009.pdf> - Acesso em 20 de maio, 2017.

COBIT - BUSINESS CONTINUITY-DISASTER RECOVERY PLANNING (ISACA) – Accesso em 20 de maio, 2017.

Commvault Books Online - <www.documentation.commvault.com> - Acesso em 19 de maio, 2017.

DELL / EMC Brasil (2017) - <https://brazil.emc.com/corporate/glossary/data-deduplication.htm> - Acesso em 20 de maio, 2017.

DOMINGUES, D. A. BACKUP E RECUPERAÇÃO MAIS EFETIVA. 2012. 36 f. Monografia (Especialização) - Curso de Software Livre Aplicado A Telemática, Eletrônica, Universidade Tecnológica Federal do Paraná, Curitiba, 2012 - Acesso em 17 de maio, 2017.

ESG Group - < http://www.esg-global.com/> - Acesso em 17 de maio, 2017.

Gartner Group - MAGIC QUADRANT FOR DATA CENTER BACKUP AND RECOVERY SOLUTIONS - <https://www.gartner.com/doc/reprints?id=1-487ZL50&ct=170731&st=sb&kui=atAWPLa52Xtd4ANq6pjNLA> – Acesso em 12 de agosto, 2017.

ISO/IEC 17799 (ABNT, 2005) – Acesso em 17 de maio, 2017.

ITIL v3 Business Continuity Book – Acesso em 17 de maio, 2017.

MORAES, E. PLANEJAMENTO DE BACKUP DE DADOS. 2007. 124 f. Dissertação (Mestrado) - Curso de Mestrado em Gestão e Desenvolvimento Regional do Departamento de Economia, Contabilidade e Administração, Universidade de Taubaté, Taubaté, 2007 - Acesso em 17 de maio, 2017.

SILVA, F. H. R. UM ESTUDO SOBRE OS BENEFÍCIOS E OS RISCOS DE SEGURANÇA NA UTILIZAÇÃO DE CLOUD COMPUTING. 2010. 15f. Artigo cientifico de conclusão de curso apresentado no Centro Universitário Augusto Motta, UNISUAM-RJ – Acesso em 20 de maio, 2017.

WILL, D. E. M. METODOLOGIA DA PESQUISA CIENTÍFICA: livro digital / Daniela Erani Monteiro Will; design instrucional / Daniela Erani Monteiro Will. – Palhoça: UnisulVirtual, 2016. 74 p. – Acesso em 12 de agosto, 2017.