



INTER
FACES
CIENTÍFICAS

EXATAS E TECNOLÓGICAS

ISSN IMPRESSO - 2359-4934

ISSN ELETRÔNICO - 2359-4942

DOI - 10.17564/2359-4934.2015v1n2p57-74

TI VERDE: UMA PROPOSTA DE ECONOMIA ENERGÉTICA USANDO VIRTUALIZAÇÃO

Luana Barreto da Silva¹
George Leite Junior³

Leonardo Henrique da Silva Bomfim²
Marcelino Nascimento de Oliveira⁴

RESUMO

Tecnologia da Informação (TI) é um dos principais responsáveis pelos problemas meio ambiente. Este é um desafio a ser superado pelos gerentes de TI para reduzir o custo ea manutenção de centros de dados. Para promover o usuário de recursos do computador em um eficiente e menos prejudiciais ao meio ambiente, o método de TI Verde propor formas sustentáveis para apoiar um datacenter. Uma dessas formas é a virtualização de datacenters, o que significa que um servidor físico tem os virtuais, trabalhando como servidor de singles. É importante analisar a viabilidade de manter um centro de dados virtualizado pela análise de disponibilidade de servidores. Se o uso da virtualização permite uma redução de custos, mas também pode tornar o sistema mais suscetível a tempo de inatividade. Este trabalho analisa a disponibilidade de dois ambientes, um com um servidor virtualizado e outro com os servidores não-virtualizados. Os serviços oferecidos são e-mail, DNS, Web Server e servidor de arquivos, um cenário típico em muitas empresas. É desenvolvido um estudo de caso

utilizando modelagem analítica com Árvore de Falhas e Cadeias de Markov. A árvore de falhas é usado para modelar os servidores e cadeias de Markov para modelar o comportamento de cada componente de hardware e software. O ambiente não virtualizado é composta de quatro servidores, cada um fornecendo serviços específicos, enquanto a virtualização consiste de um único servidor com quatro máquinas virtuais, cada um proporcionando um serviço. Ao analisar os modelos desenvolvidos, os resultados mostram que, apesar de o sistema não virtualizado tem menos tempo de inatividade, porque tem menos dependência entre os serviços, a diferença, neste caso, é de 0,06% ao ano, torna-se irrelevante quando comparado com os benefícios trazidos pela virtualização.

PALAVRAS-CHAVE

virtualização, análise de disponibilidade, cadeias de Markov, TI verde.

ABSTRACT

Information Technology (IT) is one of the main responsible for environment troubles. This is a challenge to be overcome by IT Managers to reduce the cost and the maintenance of datacenters. To promote the user of computer resources on a efficient and less harmful to environment, the Green IT method propose sustainable ways to support a datacenter. One of those ways is the datacenters virtualization, which means that one physical server has virtual ones, working as single server. It is important to analyze the viability of keep a virtualized datacenter by the analysis of availability of the servers. If the use of virtualization enables a cost reduction, it can also make the system more susceptible to downtime. This work analyzes the availability of two environments, one with a virtualized server and the other with non-virtualized servers. The services offered are e-mail, DNS, Web Server and File Server, a typical scenario in many companies. It is developed a case study using analytical modeling with Fault Tree and

RESUMEN

Tecnología de la Información (TI) es un contribuyente importante para el medio ambiente problemas. Este es un reto que hay que superar por los administradores de TI para reducir los centros de datos y mantenimiento. Para promover recursos usuario de la computadora de una manera eficiente y menos dañina para el medio ambiente, el método de Green IT proponen formas sostenibles de apoyo a un centro de datos. Una de esas formas es la virtualización de centros de datos, lo que significa que un servidor físico es virtual, trabajando como servidor de singles. Es importante analizar la viabilidad de mantener un centro de datos virtualizado para el análisis de servidores disponibles. Si el uso de la virtualización permite una reducción de costes, pero también puede hacerlo más susceptible a la inactividad del sistema. Este trabajo analiza la disponibilidad de dos habilitaciones, una con un servidor virtualizado y otro con los servidores no virtualizados. Los servicios que ofrece son el correo electrónico, DNS, servidor Web y el servidor de archivos, un escenario típico en muchas empresas.

Markov Chains. The Fault Tree is used to model the servers and Markov Chains to model the behavior of each component of hardware and software. The non-virtualized environment is composed of four servers, each one providing specific services, while the virtualized consists of a single server with four virtual machines, each one providing a service. By analyzing the models developed, the results show that although the non-virtualized system has less downtime, because has less dependence between the services, the difference in this case is 0.06% annually, becomes irrelevant when compared to the benefits brought by virtualization.

KEYWORDS

Virtualization. Analysis of Availability. Markov Chains. Green IT.

Se desarrolló un estudio de caso utilizando el modelado analítico con el árbol de fallos y cadenas de Markov. El árbol de fallos se utiliza para modelar los servidores y cadenas de Markov para modelar el comportamiento de cada componente de hardware y software. El entorno no virtualizado se compone de cuatro servidores, cada uno proporciona servicios específicos, mientras que la virtualización consiste en un único servidor con cuatro máquinas virtuales, cada uno proporciona un servicio. Mediante el análisis de los modelos desarrollados, los resultados muestran que, aunque el sistema no virtualizado tiene menos tiempo de inactividad, ya que tiene menos dependencia entre los servicios, la diferencia en este caso es de 0,06% anual, es irrelevante en comparación con los beneficios aportados por la virtualización.

PALABRAS CLAVE

virtualización, análisis de disponibilidad, cadenas de Markov, verde TI.

1 INTRODUÇÃO

A evolução tecnológica das últimas décadas tem apresentado os mais incríveis aparelhos eletrônicos, entre eles modernos celulares, televisores e computadores. Este último, tornando-se um dos campeões de demanda (KING, 2009).

Os computadores têm sido difundidos desde o início da década de 1990, e a cada ano tornam-se mais avançados e apresentando melhor desempenho (KING, 2009), atraindo a atenção do consumidor que busca substituir os equipamentos antigos por novos supercomputadores.

A substituição desses produtos gera impacto ambiental, uma vez que os equipamentos antigos são eliminados em aterros sanitários prejudicando o solo. Os novos exigem em sua construção uma grande quantidade de materiais como água, combustíveis fósseis e produtos químicos (RUTH, 2009).

Os equipamentos de Tecnologia da Informação (TI) são um dos principais responsáveis pelos problemas ambientais, desde a sua fabricação até a sua eliminação (MURUGESAN, 2010). Portanto, é preciso minimizar o impacto ambiental gerado pela TI.

É neste contexto que surge a TI Verde (AUSEKLIS, 2009), promovendo um conjunto de práticas para tornar mais sustentável e menos prejudicial o uso da computação. Dentre essas práticas está a diminuição do consumo de energia pelos aparelhos computacionais e datacenters, redução da quantidade de produtos químicos utilizados na produção dos computadores, o reuso e a reciclagem (MURUGESAN, 2008). Dessa forma é possível usufruir dos benefícios trazidos pelos avanços tecnológicos, e manter uma consciência e sustentabilidade ambiental.

A energia elétrica é uma das principais responsáveis pelas mudanças climáticas do planeta, uma vez que sua produção é dependente da queima de combustíveis fós-

seis, liberando óxido de nitrogênio, metano, dióxido de carbono e outros inúmeros poluentes (MURUGESAN, 2008). Assim, reduzir o consumo de energia elétrica é fundamental para diminuição das emissões de gases poluentes e seu impacto sobre o meio ambiente.

Com a atual crise de produção energética e o aumento das taxas cobradas pelas companhias fornecedoras de serviços de distribuição de eletricidade, torna-se um fator motivacional o estudo de técnicas que proporcionem uma redução no consumo de energia de um *datacenter*, proporcionando uma redução no valor pago mensalmente por uma empresa devido à redução no uso de quilowatts.

Uma proposta para implementar técnicas de TI Verde em um datacenter corporativo consiste em empregar a virtualização de servidores. A técnica de virtualização consiste em utilizar uma instância física de uma máquina física, e por meio desta, criar máquinas virtuais, cada uma desempenhando o papel de um servidor físico distinto, e com isto, tendo apenas o consumo de um servidor.

Este artigo discorrerá a respeito da TI Verde, dando ênfase à redução do consumo de energia pelos datacenters e explanando sobre como é possível, por meio da virtualização, reduzir este consumo. Será utilizado um estudo de caso que comprovará como a virtualização diminui despesas com energia.

Apesar dos benefícios proporcionados pela virtualização, é preciso analisar a disponibilidade de um servidor virtualizado. Uma vez que, diversos serviços são oferecidos por um único servidor físico, diferente de um datacenter tradicional, com o conceito de um serviço por servidor.

A disponibilidade é uma propriedade de dependabilidade de sistemas, que consiste na capacidade de um sistema oferecer um serviço confiável, que seja tolerante a falhas. A tolerância à falha corresponde

à capacidade do sistema do continuar provendo serviços mesmo na presença de falhas, por meio de técnicas como redundância de *hardware*, configurações de *software* e virtualização (LATH, 2010).

Para analisar a viabilidade da virtualização numa empresa, apresentando a aplicação de uma técnica de TI Verde, este trabalho emprega a modelagem analítica para avaliar a disponibilidade dos serviços, possibilitando assim que uma organização efetue a escolha de virtualizar seu *datacenter*. O estudo de viabilidade sobre implantar ou não a virtualização, em um ambiente de *datacenter*, está fundamentado na modelagem analítica, que avaliará a disponibilidade de serviços não críticos a serem migrados para o ambiente proposto nesse trabalho.

A modelagem analítica é utilizada em estudos para efetuar a avaliação de sistemas. Um exemplo está presente em Kochut (2009), que formula um modelo analítico para investigar como o consumo de energia em servidores virtuais depende de propriedades da carga de trabalho, da infraestrutura de virtualização e da média de densidade de máquinas virtuais por servidor físico. Outro exemplo é em Menasce (2005) que apresenta um modelo de performance para ambientes virtuais por meio da modelagem analítica.

Um trabalho de destaque é o de Kim (2009), que por meio da modelagem analítica apresenta um modelo de disponibilidade e análise de sistemas virtualizados. Em Kim (2009) foi construído dois sistemas, um virtualizado e outro não virtualizado, e com o uso da modelagem analítica determinaram a disponibilidade, considerando as falhas que podem ocorrer em nível de *hardware*, *software* e do próprio programa *hypervisor* que realiza a virtualização da máquina.

Desta forma, este trabalho analisa a disponibilidade de um *datacenter* que fornece os serviços de servidor de Arquivos, Web, DNS e E-mail. Estes serviços podem ser encontrados em diversas empresas sejam elas da área de tecnologia informação ou de outros setores.

A disponibilidade dos servidores é obtida com o uso da modelagem analítica, utilizando Cadeias de Markov para analisar o comportamento dos componentes de *hardware* e *software*. Com o uso de uma Árvore de Falhas foi modelado o ambiente virtualizado e um não virtualizado para posterior comparação de disponibilidade entre eles.

O artigo está organizado da seguinte forma. A seção 2 aborda sobre TI Verde. Na seção 3 é apresentado o uso de TI Verde nos *datacenters*. A seção 4 descreve o conceito de virtualização. A seção 5 discute as vantagens de um *datacenter* virtualizado para TI Verde. A seção 6 apresenta a abordagem sobre análise de disponibilidade de sistemas computacionais, abrangendo ao uso de modelagem analítica, utilizando Cadeias de Markov. Na seção 7 são discutidos os cenários analisados e na seção 8 são apresentados os modelos desenvolvidos com Árvores de Falhas e Cadeias de Markov. Na seção 9 são discutidos os resultados alcançados. E na seção 10 encontram-se as considerações finais deste trabalho.

2 TI VERDE

Atualmente, discute-se a necessidade das companhias e da sociedade criarem formas harmoniosas de relacionamento entre o avanço da tecnologia e o meio ambiente (BELTRAN, 2008). É sob este aspecto que surge a TI Verde, com o objetivo de apresentar maneiras de reduzir o consumo de energia e, conseqüentemente, diminuir as emissões de gases que contribuem para o efeito estufa (MURUGESAN, 2008), estimular a utilização de materiais menos nocivos e por fim, incentivar a reutilização e reciclagem (BELETRAN, 2008).

A redução do consumo de energia é uma ação prioritária da TI Verde. Isto ocorre porque a Tecnologia da Informação é uma das principais fontes emissoras de gases poluentes que contribuem para o aquecimento global (MURUGESAN, 2010).

Na TI o consumo de energia está na fabricação dos equipamentos computacionais, na utilização desses produtos e também na reciclagem e descarte dos mesmos (AUSEKLIS, 2009). Mas o ponto crítico está na operação, e os grandes “sugadores” atuais de energia no setor de tecnologia são os datacenters, local onde são concentrados os servidores responsáveis pelo processamento de dados de uma empresa ou organização (SKEJIC, 2010). A TI Verde ajuda a olhar para os datacenters por meio de um prisma ambiental, onde mudanças básicas podem garantir substanciais reduções no consumo de energia.

3 TI VERDE E OS DATACENTERES

A necessidade de se possuir datacenters nas organizações só tem aumentado com a crescente dependência que as empresas têm na Tecnologia da Informação (SCARFO, 2011). Não obstante, a potência dos servidores segue seu curso normal de aceleração de desempenho, e ainda, consumindo mais energia em sua operação (SKEJIC, 2010).

Proporcionalmente ao aumento do consumo total de energia pelos servidores e sistemas de refrigeração nos datacenters, amplia-se também a preocupação dos *Chiefs Information Officer* (CIOs) com a conta de energia elétrica, com o espaço físico ocupado pelas máquinas e com os custos gerados pelos aparelhos de refrigeração do ambiente (MURUGESAN, 2010).

Neste sentido, o mundo corporativo começa a adotar e, principalmente, criar ações para atender as necessidades de um negócio sustentável, conciliando consciência ecológica com práticas de negócio. As empresas Nokia, Samsung, Sony e Sony Ericsson optaram por ajustar o parque tecnológico para seguir a tendência de TI Verde, enfrentaram desafios e mudanças, e hoje possuem sucesso por meio das vantagens obtidas com o uso consciente dos recursos computacionais (WATI, 2010).

Eliminar e minimizar impactos ambientais, além de reduzir o aquecimento global, previne a ocorrência de outros desastres naturais e contribui para a melhoria da qualidade de vida do homem (MURUGESAN, 2010). Isso é motivo suficiente para que as empresas se tornem sustentáveis.

Porém, se tornar “verde”, também, contribui para a lucratividade, portanto é um incentivo que não deve ser ignorado (AUSEKLIS, 2009). Em pesquisa realizada pela Sun Microsystems australiana, com 1.500 respostas de 758 organizações de grande e pequeno porte, na Austrália e na Nova Zelândia, foi apurado que a redução do consumo de energia elétrica e a redução de custos que isso traz são as principais razões para a utilização de práticas ecologicamente corretas ou práticas “ecoresponsáveis”, seguida do menor impacto ambiental e melhoria nos sistemas (MURUGESAN, 2008).

As empresas devem desenvolver uma visão estratégica de TI Verde, criando uma política verde, delineando metas, objetivos, planos de ação e cronogramas que ajudem a reduzir os problemas ambientais (WATI, 2010). A Symantec Corp® revela que a TI Verde agora é essencial e faz parte do planejamento das empresas, segundo os próprios executivos do setor de TI. Os dados revelam que 45% dos executivos entrevistados mostram que existem iniciativas em termos de TI Verdes implementadas, principalmente para a redução do consumo energético e custos de resfriamento de equipamentos (AKANO, 2014).

No esforço conjunto dos profissionais da área de TI, a infraestrutura dos datacenters torna-se o foco de ação principal para a redução do consumo de energia e otimização da tecnologia. Uma estratégia fundamental que minimiza os gastos com eletricidade é a virtualização, que possibilita executar diversos sistemas operacionais em um único equipamento físico (MENASCE, 2004). Portanto, a necessidade de aquisição de novos equipamentos implicando em custos é extinta. Este panorama favorável à virtualização sina-

liza que há um grande mercado a ser explorado e que a grande tendência é que as empresas adotem essa tecnologia em um ritmo cada vez mais acelerado.

4 VIRTUALIZAÇÃO

O conceito de máquina virtual surgiu na década de 1960 quando a IBM desenvolveu o sistema operacional M44/44X, a partir dele, outros foram desenvolvidos com suporte a virtualização, como o OS/370 (MEYER, 2010).

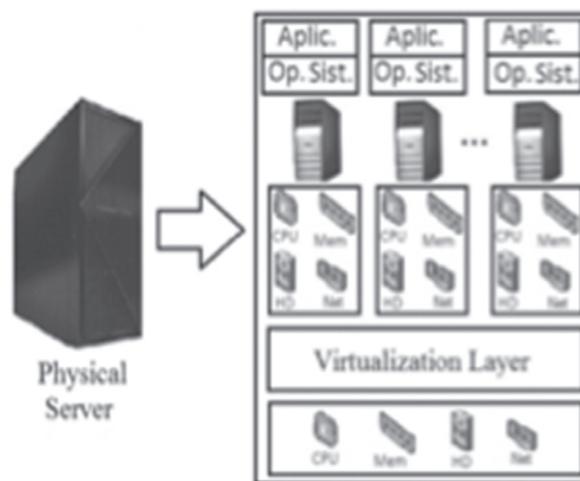
Um assunto que está em alta na área tecnológica é virtualização. Em poucas palavras pode ser explicado da seguinte maneira: vários Sistemas Operacionais compartilhando o mesmo *hardware*. Apesar de o assunto parecer novidade, essa abordagem foi usada com sucesso pela IBM no início dos anos 1960, que em uma linha de mainframes, oferecia uma máquina virtual, portada para várias de suas plataformas, sobre a qual as aplicações executavam (CAMPBELL, 2014).

A partir de 1990 o desenvolvimento do hardware torna-o com melhor desempenho e qualidade, e é desenvolvida a linguagem de programação Java que utiliza o conceito de máquina virtual para que os programas desenvolvidos sejam capazes de funcionar multiplataforma (LAUREANO, 2014).

Atualmente, a virtualização está sendo utilizada não apenas para redução de custos em *datacenters*, há também o uso na área de educação (DOBRILOVIC, 2006), teste de softwares, consolidação de servidores (MENASCE, 2004), dentre outras áreas.

A virtualização é um processo que permite executar vários sistemas operacionais em um único equipamento (RABBANI, 2012). Conforme pode ser visualizado na Figura 1, em que uma máquina física possui algumas máquinas virtuais, cada uma com recursos de *hardware*, sistema operacional e aplicações.

Figura 1 – Processo de Virtualização



Fonte:

Um ambiente de máquina virtual consiste de três partes básicas: o sistema real, que contém os recursos reais de *hardware* e *software* do sistema; o sistema virtual que executa sobre o sistema virtualizado; e a camada de virtualização, denominada *hypervisor*, que constrói as interfaces virtuais a partir da interface real (UHLIG, 2005). O *hypervisor* é um software que é executado na máquina física, destacam-se o Xen da Citrix, o Hyper-V da Microsoft e o VMWare da VMEareInc(COURCOUBETIS, 2012).

5 VANTAGENS DA VIRTUALIZAÇÃO PARA TI VERDE

A eficiência do *datacenter* é um grande problema em termos de consumo de energia. As baixas taxas de utilização dos servidores significam que as empresas estão pagando pela energia, manutenção, apoio às operações e, utilizando apenas uma pequena porcentagem da capacidade de computação de um servidor (MENASCE, 2005).

Segundo os dados Gartner, os datacenters representam cerca de 23% da energia consumida na área

de Tecnologia da Informação (RUTH, 2009). Neste aspecto a virtualização faz-se presente, promovendo um maior aproveitamento dos recursos existentes, acabando com o paradigma “um aplicativo para cada servidor”, em que os servidores ficavam subutilizados.

Com a adesão da virtualização, os *datacenters* passam por uma reestruturação em que se reduzem os custos da infraestrutura de Tecnologia da Informação. Em um ambiente de virtualização há a centralização das aplicações, retirando o máximo de aproveitamento dos recursos existentes em uma única máquina física (RUTH, 2009). Dessa forma economiza-se com aquisições futuras de hardware, refrigeração, energia elétrica e diminui a emissão de CO2 na atmosfera (SKEJIC, 2010).

É possível diminuir a infra-estrutura física dos *datacenters*, com a redução na aquisição de novos servidores, possibilitando o melhor uso do hardware e do gerenciamento por parte dos administradores (SKEJIC, 2010).

Diante das vantagens supracitadas, observa-se que a virtualização é de grande importância para atender às necessidades crescentes da computação, seja sob o aspecto da TI Verde quanto pela visão empresarial (WATI, 2010).

A tecnologia de virtualização não é só um discurso pró-ecologia. Os benefícios são concretos e comprováveis. As empresas que aderem à virtualização buscam muito mais do que a defesa da natureza. Elas procuram formas de racionalizar seus gastos, o que acaba refletindo em ganhos para o meio ambiente (AUSEKLIS, 2009).

Para que a virtualização seja adotada nas empresas, faz-se necessário um estudo de disponibilidade que garanta a oferta dos serviços. Para esse estudo são utilizadas Cadeias de Markov para investigar o comportamento de cada componente envolvido no processo de virtualização.

6 ANÁLISE DE DISPONIBILIDADE COM CADEIAS DE MARKOV

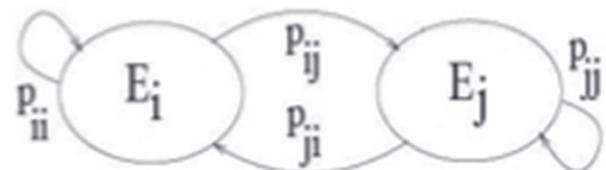
A disponibilidade de um sistema é definida como a fração de tempo em que o sistema está disponível para aceitar solicitações de serviços dos usuários. A duração de tempo em que o sistema está indisponível é chamada de tempo de indisponibilidade; e a duração de tempo em que o sistema está disponível é chamada de tempo de disponibilidade (JAIN, 1991).

Neste trabalho, para análise de disponibilidade do serviço de um servidor virtualizado é utilizado Cadeias de Markov para descrever o comportamento de cada componente de *hardware* e *software*.

Processos de Markov representam fenômenos que podem ser classificados em estados finitos e discretos, com uma probabilidade de transição entre os estados. A sequência de estados, seguindo este processo denomina-se Cadeia de Markov (DESROCHERS, 1995).

Cadeias de Markov podem ser representadas, utilizando diagramas de transições de estados, como apresentado na Figura 2. Os estados são representados pelos círculos nomeados de E_i e E_j , e as transições são p_{ij} , p_{ji} , p_{ii} e p_{jj} . A taxa total das transições de entrada e saída de um estado totalizam 1, representando a probabilidade 100%.

Figura 2 – Exemplo de Cadeia de Markov



Fonte:

As Cadeias de Markov apresentam limitações no modelo que são explicadas por (MENASCE, 2004). Uma dessas limitações é o fato dela ser *Memoryless*.

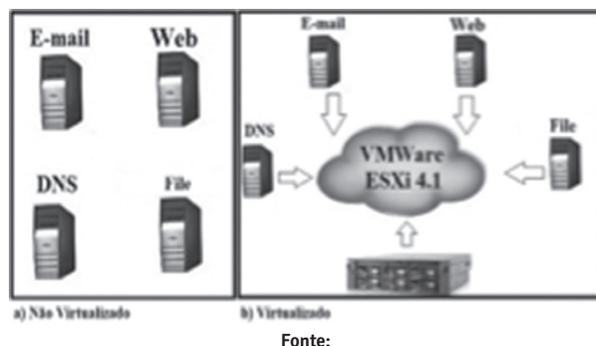
sAssumption, desta forma, assume-se que toda a informação necessária do sistema está descrito naquele estado, o que também ocasiona que é irrelevante o tempo que se passa em um mesmo estado. A única coisa importante a saber é a probabilidade de ir para um determinado estado por meio do atual.

E, outra limitação é por ser *Resulting Limitation*, porque todas as informações devem estar contidas em estados, as Cadeias de Markov estão sujeitas a terem grandes dimensões, o que causa um aumento da complexidade e perda de precisão.

7 CENÁRIOS ANALISADOS

Dois cenários são propostos para o desenvolvimento do estudo de caso. O primeiro cenário proposto é um *datacenter* não virtualizado, com o conceito de um serviço por servidor, como apresentado na Figura 3a. Com esta configuração, a paralisação de um servidor deixa apenas o serviço que ele provê indisponível, não afetando os demais.

Figura 3 – Cenários Analisados



Um *datacenter* virtualizado, com os mesmos serviços do não virtualizado, é representado na Figura 3b. Neste *datacenter* foram criadas quatro máquinas virtuais em um único servidor físico, de forma a não exceder o limite dos recursos computacionais disponíveis.

O servidor virtualizado é um modelo Blade com processador Intel Xeon E5430 descrito na Tabela 1, memória RAM de 18 GB e um disco de armazenamento de 530 GB.

Tabela 1 – Descrição do Servidor

Specification	Description
Numberof cores	4
ClockSpeed	2.66 GHz
L2 Cache	12 MB
FSB Speed	1333 MHz
Instruction Set	64-Bit

Fonte:

A virtualização do servidor é realizada com o *hypervisor* MWareESXi 4.1. A escolha deste *software* ocorre por ser uma ferramenta de licença gratuita, com documentação disponível no site do fabricante e histórico de funcionamento em outros processos de virtualização. A Tabela 2 mostra a configuração de cada máquina virtual criada.

Tabela 2 – Configuração das Máquinas Virtuais

Service	Processor	Mem	HD
Email	2x5.208 GHz	4 GB	100 GB
DNS	1x2.064 GHz	2 GB	8 GB
File	1x2.064 GHz	4 GB	100 GB
Web	1x2.064 GHz	512 MB	8 GB

Fonte:

8 MODELOS PROPOSTOS

O modelo proposto consiste em uma Árvore de Falhas para calcular a probabilidade de indisponibilidade de *hardware*, *hypervisor*, máquinas virtuais e apli-

cações, utilizando Cadeias de Markov para capturar o comportamento de cada componente do sistema.

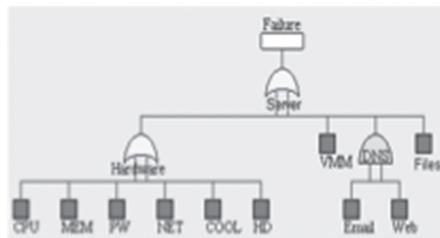
Os modelos utilizados para representar os cenários são as Árvores de Falha. Uma Árvore de Falha representa um sistema por meio de nós correspondendo às portas lógicas do tipo “OR” e do tipo “AND”.

Uma porta lógica do tipo “OR” apresenta uma falha se qualquer um dos componentes representados abaixo da porta tiver uma indisponibilidade. A porta lógica do tipo “AND” representa uma falha se todos os componentes abaixo dela apresentarem indisponibilidade.

O modelo de Árvore de Falha apresentado na Figura 4 representa o sistema virtualizado em um servidor. O primeiro nó é uma porta do tipo “OR” e corresponde ao servidor físico. Abaixo dessa porta estão a divisão de *hardware*, a camada de virtualização (VMM) e as máquinas virtuais do servidor.

A porta lógica do tipo “OR” descrita por “Hardware” apresenta os seis componentes de *hardware* do servidor (processador, memória, potência, rede, refrigeração e HD), se houver uma falha em cada um dos componentes, por ser uma porta do tipo “OR”, corresponde a uma falha no sistema.

Figura 4 – Árvore de Falha do Servidor Virtualizado



Fonte:

O servidor de DNS é representado por uma porta lógica do tipo “AND”, com as máquinas virtuais de Email e Servidor Web abaixo dela. Esta representação ocorre porque se o servidor de DNS deixar de funcio-

nar, os serviços de E-mail e Servidor Web continuam funcionando sem a tradução de endereços. Porém, caso o Servidor Web e o serviço de E-mail apresentem indisponibilidade, o serviço de DNS, mesmo estando disponível, não estará sendo utilizado.

O servidor de arquivos é representado no nível abaixo do servidor físico, pois uma falha nele ocasiona indisponibilidade do sistema, por possuir arquivos utilizados pelas demais máquinas virtuais.

Para o *datacenter* não virtualizado, a Árvore de Falha é apresentada na Figura 5. A explicação assemelha-se ao do *datacenter* virtualizado, porém, ao invés de máquinas virtuais há a representação de servidores físicos com hardware e sistema operacional.

Figura 5 – Árvore de Falha de um *datacenter* não virtualizado



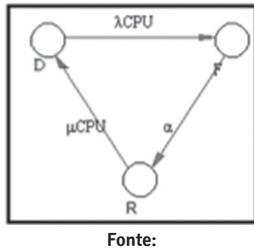
Fonte:

Os componentes de *hardware* (CPU, memória, HD, sistema de refrigeração, dispositivo de rede e fonte de energia), *software* (sistema operacional e hypervisor) e máquinas virtuais são representados por Cadeias de Markov para obtenção do estado de falha de cada um.

A Figura 6 apresenta a Cadeia de Markov para o comportamento do processador do servidor físico. No estado “Disponível” o sistema está ativo, quando o processador apresenta uma falha com uma taxa CPU, o sistema entra no estado “Falha”, e uma pessoa é de-

signada para resolver o problema com uma taxa, indo ao estado “Reparo”. Com o reparo concluído em uma taxa média de reparo μ_{CPU} , o sistema volta ao estado “Disponível”.

Figura 6 – Cadeia de Markov para o subsistema: CPU



Os subsistemas de memória, fonte de energia, sistema de refrigeração, HD e rede possuem a mesma cadeia apresentada na Figura 6, apenas com os valores dos parâmetros de entrada sendo diferentes. As Figura 7 e 8 representam, respectivamente, as Cadeias de Markov para os componentes de memória, fonte de energia, sistema de refrigeração (cooling), HD e rede.

Figure 7 – Markov Chains for the Subsystems: Memória, Energia

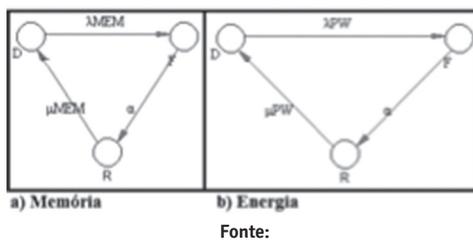
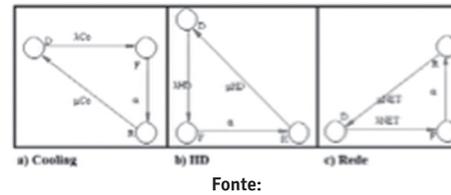


Figura 8 – Cadeia de Markov para o subsistema: Cooling, HD, Rede



A Cadeia de Markov para o comportamento do sistema operacional possui cinco estados, apresentados na Figura 9. O modelo inicia no estado “Disponível”, ocorrendo uma falha com uma taxa λ_{SO} , o modelo passa para o estado “Falha”. Após detecção da falha, a uma taxa λ_{SO} , o modelo move-se para o estado “Reinício” e o sistema é reiniciado.

Se com este procedimento o sistema voltar a funcionar, a cadeia segue para o estado “Disponível”; caso contrário, move-se para o estado “FalhaReinício”. No estado “FalhaReinício” uma pessoa é designada para corrigir o problema, indo para o estado “Reparo”. Quando o reparo for concluído, o sistema volta para “Disponível”. Neste modelo, μ_{SO} representa o tempo médio de reparo, b_{SO} fator de reinício e λ_{SO} tempo médio de reinício. A Cadeia de Markov para o *hypervisor*, apresentado na Figura 10, é semelhante ao do sistema operacional.

Figura 9 – de Markov para o subsistema: Sistema Operacional

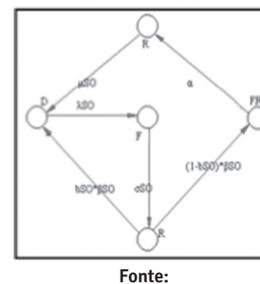
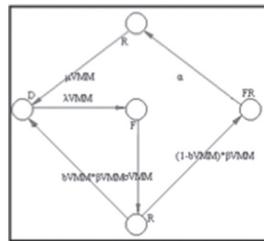


Figura 10 – de Markov para o subsistema: Hypervisor



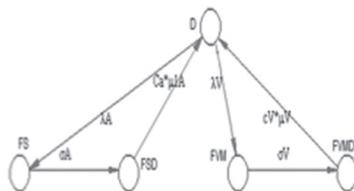
Fonte:

A Cadeia de Markov, para o comportamento das máquinas virtuais, é apresentada na Figura 11, possuindo cinco estados. No estado “Disponível” a máquina virtual está funcionando adequadamente. Ocorrendo uma falha no serviço, a cadeia passa para o estado “Falha-Servico”. Com a falha sendo detectada passa para o estado “FalhaServicoDetectada”, neste estado a falha é corrigida e a cadeia retorna para o estado “Disponível”.

Para uma indisponibilidade na Máquina Virtual, a cadeia apresentada na Figura 11, segue o mesmo comportamento da falha no serviço.

Nesse modelo, A é o tempo médio de falha de uma aplicação, A é o tempo médio para detecção da falha, CA é o fator para reparo da aplicação e μ_1A é o tempo médio de reparo da aplicação. O símbolo V é o tempo médio para falha da máquina virtual, V representa o tempo médio para detecção da falha em uma máquina virtual, CV é o fator de reparo para máquina virtual e μV é o tempo médio de reparo para máquina virtual.

Figura 11 – de Markov para o subsistema: Máquinas Virtuais



Fonte:

9 RESULTADOS E DISCUSSÕES

As Cadeias de Markov e as Árvores de Falha são modeladas, utilizando a ferramenta SHARPE. Com esta ferramenta é realizado o cálculo de disponibilidade dos componentes das Cadeias de Markov. E com as Árvores de Falha são calculadas as probabilidades de disponibilidade de cada sistema e o tempo médio de falha anual.

O SHARPE é uma ferramenta que fornece uma linguagem de especificação e métodos de solução para a maioria dos tipos de modelo utilizado para o desempenho, confiabilidade e modelagem de performance.

A ferramenta apresenta o conceito de hierarquia, porque permite que medidas de um modelo sejam utilizadas como parâmetros de entrada para outros modelos (TRIVEDI, 2002).

A obtenção dos parâmetros dos modelos é dada por meio dos manuais dos componentes, observações e trabalhos disponíveis. Os valores dos parâmetros de entrada nos modelos para o tempo médio de falha dos componentes deste estudo de caso são obtidos em (KIM, 2009).

Os valores para o tempo médio de reparo dos componentes e o tempo de um funcionário ser designado variam de acordo com o grupo de trabalho de cada empresa. Neste trabalho correspondem a uma hora o tempo médio de reparo e 30 minutos o de um funcionário ser designado.

Para realização do experimento é importante colocar todos os valores em uma mesma unidade de referência. Esta troca de unidade deve considerar a que melhor se adequar ao estudo, produzindo menor número de casas decimais, e desta forma, reduzindo o erro nas operações aritméticas.

Os valores dos parâmetros, deste estudo de caso, são apresentados na Tabela 3.

Tabela 3 – Valor dos parâmetros dos modelos

Parâmetro	Descrição	Valor
1/λCPU	Tempo médio de falha daCPU	2.500.000 horas
1/ λMEM	Tempo médio de falha da Memória	480.000 horas
1/ λPW	Tempo médio de falha do Sistema de energia	670.000 horas
1/ λNET	Tempo médio de falha da Rede	120.000 horas
1/ λCO	Tempo médio de falha ddcooling	3.100.000 horas
1/ λHD	Tempo médio de falha da HD	20.000.000horas
1/ λVMM	Tempo médio de falha daHypervisor	2880 horas
1/ λSO	Tempo médio de falha do Sistema Opercional	1440 horas
1/ λV	Tempo médio de falha da Máquina Virtual	2880 horas
1/ λA	Tempo médio de falha do software	336 horas
1/μCPU	Tempo médio de reparo daCPU	1 hora
1/ μMEM	Tempo médio de reparo da Memória	1 hora
1/ μPW	Tempo médio de reparo do Sistema de energia	1 hora
1/ μNET	Tempo médio de reparo da Rede	1 hora
1/ μCO	Tempo médio de reparo ddcooling	1 hora
1/ μHD	Tempo médio de reparo do HD	1 hora
1/ μVMM	Tempo médio de reparo doHypervisor	1 hora
1/ μSO	Tempo médio de reparo do Sistema Operacional	1 hora
1/ μV	Tempo médio de reparo da Máquina Virtual	1 hora
1/ μ1A	Tempo médio de reparo do Software	1 hora
1/σVMM	Tempo médio para detector a falha no hypervisor	30 segundos
1/ σSO	Tempo médio para detector a falha no sistema operacional	30 segundos
1/ σV	Tempo médio para detector a falha na máquina virtual	30 segundos
1/ σA	Tempo médio para detector a falha do software	30 segundos
1/βVMM	Tempo médio para reiniciar ohypervisor	10 minutos
1/ βSO	Tempo médio para reiniciar o Sistema operacional	10 minutos
1/α	Tempo médio para um usuário ser designado para realizar reparo	30 minutos

Parâmetro	Descrição	Valor
1/bVMM	Coeficiente de reinício doHypervisor	0.9
1/bSO	Coeficiente de reinício do Sistema Operacional	0.9
1/CV	Coeficiente de reparo da máquina virtual	0.95
1/CA	Coeficiente de reparo do software	0.9

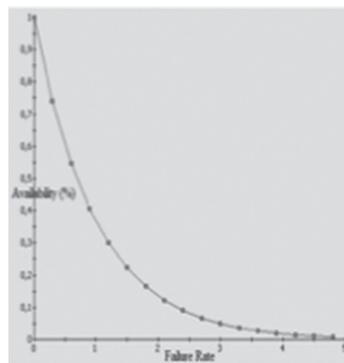
Fonte:

Desta forma, é realizada a análise do modelo para averiguar o comportamento das cadeias e árvores de falha propostos com a variação dos parâmetros. Também é realizado o cálculo da disponibilidade dos sistemas e a análise transiente. Com a análise transiente observa-se o comportamento dos sistemas com o aumento do tempo de funcionamento de ambos.

A. Análise do Modelo

A Figura 12 apresenta a análise da Árvore de Falha do sistema virtualizado. Quando há um aumento na taxa de falha da CPU, a disponibilidade do sistema decresce. Como a Árvore de Falha do sistema não virtualizado é desenvolvido, utilizando a mesma estrutura, também apresenta o mesmo comportamento.

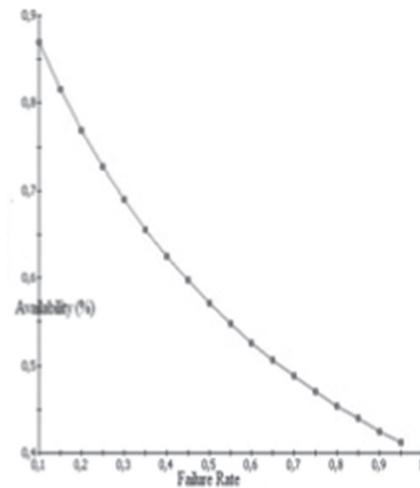
Figura 12 – Análise de disponibilidade do modelo virtualizado



Fonte:

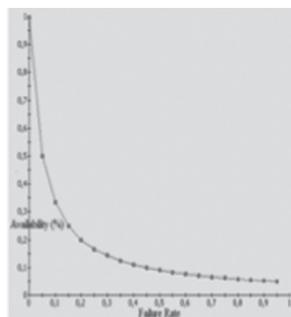
O comportamento das Cadeias de Markov para os dispositivos de *hardware* e *software* estão representados na Figura 13, com a análise específica da cadeia referente ao dispositivo físico de memória. Quando a taxa de falha aumenta há o decréscimo da disponibilidade do sistema. Comportamento semelhante ocorre com a Cadeia de Markov para as máquinas virtuais, conforme apresentado na Figura 14.

Figura 13 – Análise de disponibilidade utilizando Cadeias de Markov para os subsistemas de hardware



Fonte:

Figura 14 – Análise de disponibilidade utilizando Cadeias de Markov para o subsistema de Máquina Virtual



Fonte:

Esta análise demonstra que os modelos estão apresentando comportamento coerente com variações nos parâmetros fornecidos, ou seja, com o aumento da taxa de falha a disponibilidade decresce em todos os modelos.

B. Disponibilidade dos Sistemas

A Tabela 4 apresenta os resultados para disponibilidade e o tempo médio anual de indisponibilidade dos sistemas.

Tabela 4 – Resultado de disponibilidade dos sistemas

	Measures	Values
Sistema Virtualizado	Probabilidade de Disponibilidade	99,92%
	Tempo médio anual de indisponibilidade	468 minutos
Sistema Não Virtualizado	Probabilidade de Disponibilidade	99,98%
	Tempo médio anual de indisponibilidade	129 minutos

Fonte:

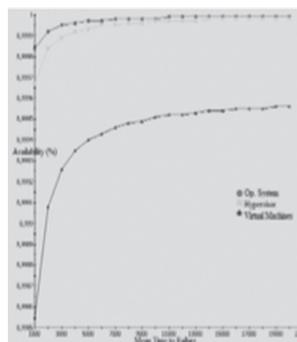
A disponibilidade do servidor virtualizado corresponde a aproximadamente 99,92%; o tempo de indisponibilidade por ano é de aproximadamente 468 minutos (7 horas e 48 minutos), o que representa 0,09% do ano.

Para o sistema não virtualizado, representado com quatro servidores físicos distintos, a disponibilidade do sistema corresponde a aproximadamente 99,98%; e tempo de indisponibilidade por ano é de aproximadamente 129 minutos (2 horas e 9 minutos), o que representa 0,03% do ano.

Essa diferença de disponibilidade entre os ambientes virtualizados e não virtualizados pode ser visualizado no gráfico apresentado na Figura 15, que corresponde à análise transiente dos sistemas. O eixo X correspondendo ao tempo médio de falha do siste-

ma operacional, *hypervisor* e máquinas virtuais em horas e o eixo Y a disponibilidade.

Figura 15 – Análise transiente dos modelos propostos



Fonte:

No ambiente virtualizado há apenas um servidor para todos os serviços fornecidos. Com a indisponibilidade do servidor físico, todos os quatro serviços virtualizados param de operar. Enquanto com quatro servidores distintos, se apenas um tiver indisponível, apenas um serviço deixa de ser fornecido.

O importante não é apenas obter os resultados, mas analisá-los. A diferença de disponibilidade entre os datacenters correspondem a aproximadamente 0,06% anual, o que para muitas empresas que prestam este serviço tal diferença torna-se irrelevante, quando comparada às vantagens trazidas pela virtualização.

10 CONCLUSÕES

Os problemas ambientais estão em destaque nos dias atuais. As empresas buscam formas de reduzir a poluição produzida, e ao mesmo tempo, diminuir os custos internos no processo de produção.

A proposta de TI Verde fornece as empresas um ideal para redução dos custos com consumo energético, por meio da virtualização de servidores.

A virtualização torna-se um procedimento adequado para este propósito, por consolidar vários servidores em uma quantidade menor, trazendo os benefícios citados.

O processo de virtualização está cada vez mais presente nos datacenters das empresas. Os lucros que se obtém com o uso deste conceito abrangem desde economia com aquisição de novos servidores físicos que ficam com recursos subutilizados, até custos com despesas de energia elétrica na empresa.

Para adoção de uma nova tecnologia é necessário realizar um estudo comparativo baseado em quais serviços a empresa disponibiliza, e desta forma garantir a viabilidade de um *datacenter* virtualizado. Este trabalho apresenta este comparativo para uma

empresa que possui serviços com o de E-mail, Servidor Web, Servidor de Arquivos e DNS.

Os resultados obtidos para análise de disponibilidade mostram que o sistema não virtualizado apresenta uma menor indisponibilidade por ter menor dependência entre os serviços, ou seja, por cada serviço estar em uma máquina diferente.

Entretanto, a diferença apresentada não foi elevada, considerando o percentual de minutos de indisponibilidade por ano abaixo de 0,1%.

Os ganhos obtidos com a virtualização transpõem o percentual de indisponibilidade para as empresas que não apresentam os serviços analisados como sendo críticos.

A economia com o menor gasto na compra de servidores, menor custo no consumo energético das empresas, ganhos com gerenciamento centralizado dos servidores e uma política ambiental pela menor emissão de gases poluentes, devem ser pontos prioritários na gestão de qualquer empresa.

REFERÊNCIAS

AKANO, A.K. A Cross-Cultural Survey of the Impact of Organizational Culture on Adoption of Green IT. **Eighth International Conference on Complex, Intelligent and Software Intensive Systems (CISIS)**, Birmingham, 2014. p.177-184.

AUSEKLIS, N. and HARMON, R. **Sustainable IT services: Assessing the impact of green computing practices**. Portland International Conference on Management of Engineering & Technology. Portland, USA, 2009. p.107-1717.

AVIZIENIS A.; LAPRIE, J.; RANDELL, B. and LANDWEHR C. **Basic concepts and taxonomy of dependable and secure computing**. IEEE Transactions on Dependable and Secure Computing, 2004. p.11-33.

BELTRAN, V.; et al. **Reducing wasted resources to help achieve green data centers.** IEEE International Symposium on Parallel and Distributed Processing, 2008. Miami, USA, 2008. p.1-8.

CAMPBELL, W.M., **Cultural Transformation to Support the Adoption of Green IT.** 28th International Conference on Advanced Information Networking and Applications Workshops (WAINA), Victoria, BC, 2014. p.554-559.

COURCOUBETIS, C. **Economic Issues in Shared Infrastructures.** IEEE/ACM Transactions on Networking, v. 20, 2012. p.594-608.

DESROCHERS, A.; AL'JAAR, R. Y. **Applications of petri nets in manufacturing systems:** Modelling, Control and Performance Analysis. IEEE Press, 1995.

DOBRILOVIC, D.; STOJANOV, Z. Virtualization software in operating systems course. ITRE 06 - **Information Technology:** Research and Education, 2006, p.222-226.

JAIN, R. **The Art Of Computer Systems Performance Analysis.** [S.l.B]: John Wiley e Sons Inc, 1991.

KIM, D. S.; MACHIDA, F.; TRIVEDI, K. S. **Availability modeling and analysis of a virtualized system.** IEEE Proceedings of the IEEE International Symposium Pacific Rim Dependable Computing, 2009. p.365-371.

KING, W.K. AND LAND, S.K. A Historical Perspective of the IEEE Computer Society: Six Decades of Growth with the Technology It Represents. **IEEE Conference on the History of Technical Societies,** 2009. p.1-10.

KOCHUt, A. Power and performance modeling of virtualized desktop systems. **IEEE International Symposium on Modeling, Analysis e Simulation of Computer and Telecommunication Systems,** 2009. p.1-10.

LATH, R.; MOHAPATRA, S. and SAHOO, J. Virtualization: A Survey on Concepts, Taxonomy And Associated Security Issues. **Second International Conference on Computer and Network Technology.** Bangkok, Thailand, 2010. p.222-226.

LAUREANO M. A. P.; MAZIERO, C. A. Virtualização: conceitos e aplicações em segurança. [S.l.B]: **Anais do 26 Simpósio Brasileiro de Segurança da Informação,** 2014.

MENASCE, D. Virtualization: Concepts, applications, and performance modeling. **CMG-Conference,** 2005. p.407-417.

MENASCE, D.; DOWDY, L.; ALMEIDA, V. **Performance by design:** computer capacity planning by example. [S.l.B]: Prentice Hall, 2004.

MEYER, R.; SEAWRIGHT, L. A virtual machine time-sharing system. **IBM Systems Journal,** v.9, 1970. p.199-218.

MURUGESAN, S. Making IT Green. **IT Professional,** v.12, 2010. p.4-14.

MURUGESAN, S. Harnessing Green IT: Principles and Practices. **IT Professional. Piscataway,** USA, 2008. p.24-33.

RABBANI, M. et al. Data Center Network Virtualization: A Survey. **IEEE Communications Surveys & Tutorials,** 2012. p.1-20.

RUTH, S. Green IT More Than a Three Percent Solution? **IEEE Internet Computing.** Piscataway, USA, 2009. p.74-78.

SCARFO, A. The Evolution of Data Center Networking Technologies. **First International Conference on Data Compression, Communications and Processing (CCP),** 2011. p.172-176.

SKEJIC, E.; DZINDO, O.; DEMIROVIC, D. Virtualization of hardware resources as a method of power savings in data center. **Proceedings of the 33rd International Convention**. Opatija, Croácia, 2010. p.636-640.

TRIVEDI, K. **SHARPE 2002**: Symbolic Hierarchical Automated Reliability and Performance.

UHLIG, R. et al. Intel virtualization technology. **Computer**, v.38, 2005, p.48-56.

WATI, Y. The Green IT Practices of Nokia, Samsung, Sony, and Sony Ericsson: Content Analysis Approach. 43rd Hawaii, **International Conference on System Sciences**, Hawaii, 2010. p.1-10.

Recebido em: 7 de Março de 2015
Avaliado em: 10 de Março de 2015
Aceito em: 2 de Abril de 2015

1. Mestre em Computação (UFS), Coordenadoria de Informática, Instituto Federal de Sergipe. E-mail: luana.silva@ifs.edu.br
2. Mestrando em Computação (UFS), Departamento de Computação, Universidade Federal de Sergipe. E-mail: leonardohsb@gmail.com
3. Coordenadoria de Informática, Instituto Federal de Sergipe. E-mail: georgeleitejunior@gmail.com
4. Departamento de Computação, Universidade Federal de Sergipe. E-mail: marcelinolive@gmail.com

