GUILHERME CARVALHO JANUÁRIO

Método para Avaliação de Sistema de Gerenciamento de Redes
Orientado a Eficiência Energética

Dissertação apresentada à Escola
Politécnica da Universidade de São
Paulo para obtenção do título de
Mestre em Ciências

São Paulo
2014

GUILHERME CARVALHO JANUÁRIO

Método para Avaliação de Sistema de Gerenciamento de Redes
Orientado a Eficiência Energética

Dissertação apresentada à Escola
Politécnica da Universidade de São
Paulo para obtenção do título de
Mestre em Engenharia

Área de Concentração:
Engenharia da Computação

Orientadora:
Prof^a. Dr^a. Tereza Cristina Melo de Brito Carvalho

São Paulo
2014

**Este exemplar foi revisado e corrigido em relação à versão original, sob
responsabilidade única do autor e com a anuência de seu orientador.**

São Paulo, de abril de 2014.

Assinatura do autor _____

Assinatura do orientador _____

Catálogo-na-publicação

Januário, Guilherme Carvalho

**Método para avaliação de sistema de gerenciamento de
redes orientado a eficiência energética / G.C. Januário. -- versão
corr. -- São Paulo, 2014.**

140 p.

**Dissertação (Mestrado) - Escola Politécnica da Universidade
de São Paulo. Departamento de Engenharia de Computação e
Sistemas Digitais.**

**1.Redes de computadores (Gerenciamento) 2.Energia (Efi-
ciência) 3.Sustentabilidade 4.Ambiente de experimentação
5.Disponibilidade I.Universidade de São Paulo. Escola Politéc-
nica. Departamento de Engenharia de Computação e Sistemas
Digitais II.t.**

AGRADECIMENTOS

A meus pais Renata e Tarcísio, pela estrutura e apoio oferecidos durante meus estudos e antes.

A minha orientadora, Professora Tereza Cristina Melo de Brito Carvalho, pela orientação, oportunidade de estudos e introdução mais profunda na área de pesquisa, acompanhadas por incentivo de desenvolvimento profissional e pessoal.

Aos professores Marco Túlio e J. J. Neto pelo apoio nos momentos iniciais de minha vida de pós-graduando.

Aos colegas de laboratório e pesquisa Carlos H. A. Costa e Marcelo C. Amaral, por ajudarem-me na ambientação da pesquisa e por aceitarem-me tão bem num projeto já em andamento, oferecendo oportunidades de expansão de trabalhos por eles previamente concebidos. Muitas das referências utilizadas aqui me chegaram por seu intermédio.

Aos colegas Ana Carolina Riekstin e Bruno Bastos Rodrigues, pela contribuição nos momentos finais de elaboração desta dissertação.

A Rafael, Lillyane, Camila, Enlai e Alan pela convivência, e a Leo e Koga também pelas discussões epistemológicas e acadêmicas e pela parceria em outros projetos de pesquisa e desenvolvimento. Estes dois muito me incentivaram nos estudos e oportunidades.

À Ericsson Telecomunicações do Brasil, pelo suporte financeiro, e a Catalin Meirosu, da Ericsson *Research* da Suécia, pelas contribuições durante a pesquisa.

À Fundação para o Desenvolvimento Tecnológico da Engenharia por mediar o suporte financeiro concedendo-me uma bolsa de estudos.

Ao LASSU e ao PCS, por me terem aceitado como mestrando.

Finalmente, agradeço à Universidade de São Paulo e seus contribuintes.

“[...]A Lei deveria ser acessível a todos e sempre, pensava ele [...]”

(Em frente à Lei (Vor dem Gesetz), F. Kafka)

“[...] O respeito à idade parece basear-se no fato de que a honra das pessoas jovens, embora admitida por pressuposição, ainda não foi colocada à prova; por conseguinte, consiste propriamente num crédito. [...]”

(Aforismos para a sabedoria de vida, Schopenhauer.
Editora: Martis Fontes. Tradução: Jair Barboza)

RESUMO

Funcionalidades de eficiência energética vêm sendo integradas a protocolos e sistemas de gerenciamento de redes. Várias dessas funcionalidades afetam a rede de diferentes modos, acarretando compromissos. Podem afetar a confiabilidade e a disponibilidade da rede ao colocarem equipamentos e enlaces em modo dormente; e podem afetar o tempo de vida dos equipamentos ao alterarem seu modo de utilização. Os resultados da integração dessas funcionalidades à rede podem ser avaliados por emulação, que provê informação sobre como algum mecanismo de roteamento em particular, fruto dessa integração, agiria em diferentes condições de rede. Preparar um ambiente de experimentação que trate das interações entre diferentes funções da rede, considerando as funcionalidades de eficiência energética, traz alguns desafios. Uma combinação de emulação e implementação das principais funcionalidades de eficiência energética provê uma visão mais próxima sobre o que pode ocorrer em implantações reais de sistemas de gerenciamento energeticamente eficiente. Esta dissertação discute (i) como se avaliam os compromissos existentes no emprego de técnicas de eficiência energética, mostrando também como o tempo de vida de equipamentos pode ser alterado e como a confiabilidade e disponibilidade da rede podem ser degradadas, e discute (ii) requisitos para o desenvolvimento de um ambiente de experimentação para avaliação de tais técnicas. O ambiente é baseado em roteadores implementados por *software*, para emulação de funcionalidades ainda não disponíveis, mas desejadas, de equipamentos de rede. Após a experimentação do protótipo de um sistema de gerenciamento de redes orientado por política de eficiência energética, aplicam-se os métodos de avaliação dos compromissos ocasionados por este sistema.

Palavras-chave: Gerenciamento de Redes de Computadores, Eficiência Energética, Sustentabilidade em Sistemas de Tecnologia da Informação e comunicação, Ambiente de Experimentação, Disponibilidade em Redes.

ABSTRACT

Energy efficiency features are being integrated in network protocols and management systems. Many of such features affect the network in different ways, thus yielding tradeoffs. They can affect the reliability and availability of the network when they put devices or links into sleep mode, and they can affect the lifetime of the devices as a result of the new operating pattern. The effects of integrating these features to the network can be assessed through emulation, which can provide insight on how a particular routing mechanism would perform in different network conditions. In addition, building an experimentation environment that is able to comprehensively account for interactions between different network functions taking into account energy efficiency features is challenging. A combination of emulation and implementation of major energy efficiency features provides a view closer to what may happen in a real deployment of energy-efficient management systems. This dissertation discusses (i) how to evaluate the tradeoffs existing between the use of energy-efficiency techniques and the lifetime of devices and with the reliability and availability of the network, and (ii) requirements for the development of a testbed to evaluate such techniques. The testbed is based on software routers, so to emulate future and desired network equipment functionalities. After experiments with the prototype of a sustainability-oriented network management system based on policies, an evaluation about the tradeoffs brought by such system to the network is performed.

Keywords: Computer Networks Management, Energy Efficiency, Sustainability in Information and Communication Technology, Testbed, Availability in Networking.

LISTA DE ILUSTRAÇÕES

Figura 1- Curva da banheira, mostrando a evolução da taxa de falha de um dispositivo	46
Figura 2 - MTTF relativo a um núcleo de processador de tecnologia 95-nm, para diferentes mecanismos de falha, quando diferentes ganhos de energia são obtidos (Rosing, 2007)	69
Figura 3 - Modelos de Márkov para confiabilidade e disponibilidade usados pelo REASoN, com taxa de cobertura de 100%. Título original: <i>Extended Markov model of a router with two connections in cold standby that considers the time to activate the standby</i> (Amaral, 2012).....	75
Figura 4 - Algumas disposições que permitem elementos (enlaces ou dispositivos) redundantes	81
Figura 5 - Modelos de Márkov para cálculo de disponibilidade para as disposições da Figura 4.....	82
Figura 6 - Arquitetura do SustNMS (Costa, 2012a)	94
Figura 7 - Requisitos para ambiente de experimentação voltado à avaliação de funcionalidades de sistemas de gerenciamento energeticamente eficiente	104
Figura 8 - Topologia integrada ao ambiente de experimentação	111
Figura 9 - Perfil de tráfego aplicado ao ambiente de experimentação proposto ..	112
Figura 10 - Consumo de energia de cada caminho em diferentes condições de tráfego.....	115
Figura 11 - Mudanças aplicadas pelo SustNMS à topologia. Taxa de amostragem de 1 min.	117
Figura 12 - Consumo de energia de cada experimento	117
Figura 13 - Consumo de energia agregado por taxa de amostragem, com ganhos de energia no destaque. Experimento (iii).	118
Figura 14 - Sobrecusto devido a pacotes de controle e monitoramento, durante 70 minutos, para os experimentos (ii), (iii) e (iv), alocado por taxa de amostragem .	120

Figura 15 - Sobrecusto dividido pela redução de consumo, alocado por intervalo entre amostragens. Experimento (ii)121

Figura 16 - Avaliação da confiabilidade para cada mudança do estado da rede, nas situações de: 1) todos os dispositivos ligados; 2) apenas roteador 3 dormindo; e 3) apenas roteador 5 dormindo.....123

LISTA DE QUADROS

Quadro 1 - Comparação entre os métodos REASoN e Reasoning	79
Quadro 2 - Quantidade diária de cada ciclo térmico, por roteador, por tipo de operação	124

LISTA DE ABREVIATURAS E SIGLAS

- ACPI** Interface de Potência e Configuração Avançada (*Advanced Configuration and Power Interface*)
- ALR** Taxa Adaptativa de Enlace (*Adaptive Link Rate*)
- ATM** *Asynchronous Transfer Mode*
- AtoM** Qualquer Transporte sobre MPLS (*Any Transport over MPLS*)
- bit** dígito binário (*binary digit*. Atualmente dicionarizada em inglês e português)
- CLI** Interface de Linha de Comando (command-line (user) interface)
- FEC** Classe de Equivalência de Encaminhamento (*Forwarding Equivalence Class*)
- FFR** Rerroteamento Rápido (*Fast ReRouting*)
- IETF** Força-Tarefa de Engenharia na *Internet* (*Internet Engineering Task Force*)
- IP** Protocolo de *Internet* (*Internet Protocol*)
- LSR** Roteador de Troca de Rótulo (*Label Switch Router*)
- LTC** Controlador de Tráfego do Linux (*Linux Traffic Controller*)
- MCTF** Número Médio de Ciclos até a Falha (*Mean Cycles to Failure*)
- MIB** Base de Informações de Gerenciamento (*Management Information Base*)
- MPLS** Comutação Multiprotocolo de Rótulos (*Multiprotocol Label Switching*)
- MTTF** Tempo Médio até a Falha (*Mean Time to Failure*)
- NIC** Placa de Rede (*network interface controller / card*)
- OSI** Interconexão Aberta de Sistemas (*Open Systems Interconnection*)

OSPF Abrir o Caminho Mais Curto Primeiro (*Open Shortest Path First*)

RFC Pedido para Comentário (*Request for Comments*)

SNMP Protocolo Simples para Gerenciamento de Rede (*Simple Network Management Protocol*)

SOFR Soma de Taxa de Falhas (*Sum-Of-Failure Rates*)

SSC Coalescimento Sincronizado no Comutador (*Synchronized Switch Coalescing*)

SSH Casca segura (*secure shell*), um protocolo para conexão segura entre dois computadores

TDDb ruptura dielétrica dependente do tempo (*time dependent dielectric breakdown*)

TIC Tecnologias da Informação e Comunicação

VPN Rede Privada Virtual (*Virtual Private Network*)

LISTA DE SÍMBOLOS

t	denota um instante de tempo
t_0	instante inicial dum intervalo de tempo
t_1	algum instante num intervalo de tempo
$R(t)$	confiabilidade relativa a t (cf. o símbolo R , abaixo)
$P(A > a)$	probabilidade do valor real da grandeza A ser maior que a
$F(t)$	(função de) probabilidade de dado evento ter ocorrido até t (nesta obra, o evento é uma falha, a não ser que diferentemente explicitado)
$f(t)$	função densidade de probabilidade de dado evento ocorrer em t
$Z(t)$	taxa instantânea de falhas relativa a t
λ	[lambda] denota uma taxa de falhas. É usada em lugar de $Z(t)$ quando a taxa instantânea de falhas é constante no tempo
e	número de Euler (aproximadamente 2,71828)
α	[alfa] denota o parâmetro de escala de uma distribuição de Weibull
a^b	denota o valor a elevado à b -ésima potência
β	[beta] denota o parâmetro de forma de uma distribuição de Weibull
C	conjunto conector
D	conjunto desconector
J	densidade de corrente
E_a	energia de ativação de um mecanismo
k	constante de Boltzmann
K	kelvin (medida de temperatura)
eV	elétron-volt (medida de energia)
N	número médio de ciclos (térmicos) até a falha (MCTF)
q	constante de exponenciação numa relação de Coffin-Manson
E	denota um enlace

- R denota um roteador (cf. o símbolo $R(t)$, acima)
- S denota um estado dum modelo de Márkov
- A matriz que relaciona as probabilidades dum sistema estar em diferentes estados em diferentes momentos
- $O(.)$ (*big-o*) denota a taxa de crescimento, ou complexidade, de uma grandeza
- I carga de trabalho em um roteador

SUMÁRIO

1	Introdução	19
1.1	Motivação	20
1.2	Objetivo.....	21
1.3	Organização	23
2	Método	25
2.1	Do ambiente de experimentação para sistemas de gerenciamento energeticamente eficiente de redes.....	25
2.2	Do método de avaliação de confiabilidade e disponibilidade	26
2.3	Do método de avaliação do tempo de vida dos dispositivos.....	26
2.4	Considerações do capítulo	28
3	Trabalhos relacionados e embasamento teórico	29
3.1	Sistemas de gerenciamento voltados a eficiência energética.....	29
3.1.1	Green OSPF	31
3.1.2	ElasticTree	32
3.2	Funcionalidades de equipamentos voltadas a eficiência energética	33
3.2.1	ACPI.....	35
3.2.2	Taxa Adaptativa de Enlace	36
3.2.3	Coalescimento sincronizado no comutador.....	37
3.3	Ambientes de experimentação	38
3.4	Gerenciamento orientado por política	39
3.4.1	Aspectos do gerenciamento orientado por política.....	41
3.4.2	Contínuo de Política	42
3.5	O protocolo MPLS	43
3.6	Conceito de tempo de vida e a distribuição de Weibull	45

3.7	Modelo de Márkov	47
3.8	Confiabilidade, disponibilidade e taxa de falhas	48
3.8.1	Cálculo de confiabilidade e disponibilidade para redes sustentáveis	50
3.9	Método dos conjuntos conectores ou dos conjuntos desconectores	51
3.10	Considerações do capítulo	53
4	Método para avaliação do tempo de vida de equipamentos	55
4.1	Visão geral sobre a questão do tempo de vida	55
4.2	Tempo de vida dos processadores	58
4.2.1	Eletromigração	59
4.2.2	Migração por estresse	61
4.2.3	Ruptura dielétrica dependente do tempo	62
4.2.4	Ciclos térmicos	62
4.3	Tempo de vida das ventoinhas	64
4.3.1	Tempo de vida do rolamento	65
4.3.2	Tempo de vida da bobina	66
4.3.3	Ciclos térmicos	67
4.4	Método de análise	68
4.5	Considerações do capítulo	70
5	Método para avaliação da confiabilidade e disponibilidade de redes sustentáveis	72
5.1	Contexto do método proposto	72
5.2	O método REASoN	73
5.3	Método proposto – do REASoN ao Reasoning	78
5.3.1	Solução analítica	83
5.3.2	Algoritmo iterativo	84

5.4	Considerações do capítulo	89
6	Arquitetura e implementação de um sistema para eficiência energética.....	91
6.1	Visão geral do sistema SustNMS	91
6.2	Arquitetura	93
6.3	Descrição dos submódulos da arquitetura do SustNMS.....	95
6.4	Detalhes da implementação	99
6.5	Considerações do capítulo	101
7	Ambiente de experimentação para soluções sustentáveis.....	103
7.1	Requisitos para o ambiente de experimentação	103
7.2	Implementação do ambiente de experimentação	106
7.3	Considerações do capítulo	108
8	Descrição dos experimentos e seus resultados	109
8.1	Políticas de sustentabilidade	109
8.2	Topologia da rede.....	110
8.3	Perfil de tráfego	111
8.4	Perfis de consumo	112
8.5	Ganhos de energia	114
8.6	Sobrecusto de operação do sistema	119
8.7	Confiabilidade	121
8.8	Tempo de vida	124
8.9	Considerações do capítulo	125
9	Considerações finais	127
9.1	Contribuições.....	129
9.1.1	Publicações.....	129
9.2	Trabalhos Futuros.....	130

REFERÊNCIAS	132
-------------------	-----

1 Introdução

Este trabalho disserta no contexto da aplicação de sustentabilidade em redes de computadores. Mais especificamente, ele discute tecnologias de rede focadas em eficiência energética. Quais são, como testar e como medir e aferir seus efeitos sobre a rede são as questões que norteiam esta dissertação. Responder à primeira questão define a natureza de um sistema de gerenciamento de redes orientado a eficiência energética. Responder às demais define um método para avaliar seus impactos.

Relativamente a outros setores da atividade humana, redes de computadores gastam muita energia elétrica e emitem muito dióxido de carbono (GeSI, 2012). São responsáveis mesmo pela emissão de outros gases em quantia equivalente a muito dióxido de carbono. A emissão excessiva não é apenas relativa a outros setores da atividade humana, mas absoluta, pois redes de computadores são, em geral, superdimensionadas (Bolla, 2011b). Não só uma rede costuma ser superdimensionada (Zhang, 2010), mas ainda seus elementos constituintes, como roteadores e comutadores, não são projetados para serem energeticamente eficientes (Bolla, 2011a).

Logo, redes de computadores abarcam um campo profícuo onde se explorarem eficiência energética e sustentabilidade. Diversos trabalhos citam técnicas verdes, isto é, que de algum modo melhoram a sustentabilidade, para redes de computadores. Neste trabalho é explorado como testar tais técnicas, como estas podem ser mais bem aproveitadas, quais seus efeitos e como medi-los. As seções seguintes continuam a contextualizar este trabalho, descrevendo sua motivação e objetivos.

1.1 Motivação

A *Global e-Sustainability Initiative* (GeSI, 2008) relatou em 2008 que o setor de TIC (Tecnologia da Informação e Comunicação) (incorporado no relatório pelos dispositivos dos usuários finais, pelas redes de telecomunicação e pelos *data centers*) foi responsável por 2% das emissões de gases do efeito estufa daquele ano (GeSI, 2008), um nível semelhante ao da aviação civil. Para tal quantia, esperava-se aumento de 6% ao ano, até 2020. Em 2012, um novo relatório (GeSI, 2012) asseverou que o aumento até então tinha de fato sido de 6.1% ao ano, mas ajustou a nova expectativa para 3.8% ao ano, até 2020. Tal quantia se deve ao exacerbado aumento do tráfego global, para o qual Lange (2011) projeta aumento exponencial até 2015.

O aumento de tráfego é principalmente alavancado pelo aumento do número de clientes de serviços de rede e pelo aumento da banda fornecida, aumento o qual incentiva os clientes a utilizarem mais intensamente os serviços adquiridos. Projeta-se que haverá mais de dois bilhões de usuários de Internet em 2015 e que, em 2020, eles excederão três bilhões (Sorensen, 2009). Justamente para abarcar essa demanda, Bolla (2011b) diz que os provedores de serviço de rede vêm anunciando planos para aumentar sua infraestrutura.

Junte-se a esse fato que provedores de serviços de telecomunicação superdimensionam a largura de banda e utilizam redundância de equipamentos, visando a assegurar o cumprimento dos acordos de nível de serviço (Zhang, 2010). O resultado é os enlaces das redes de transporte serem tipicamente utilizados muito abaixo de sua capacidade. Ora, operar os nós a capacidade máxima, a todo momento e a despeito do tráfego demandado e de sua distribuição caracteriza ineficiência energética na infraestrutura da rede. Deste modo, pode-se inferir que há grandes oportunidades para se reduzir o consumo energético de sistemas de TIC e de serviços prestados por esses sistemas.

Nesse contexto, os provedores de serviço de rede passam a se preocupar simultaneamente com o aumento do consumo de energia dispendido na operação da rede (Bolla, 2011b), (GeSI, 2008). A dupla preocupação tem também fundamento econômico, já que, mediante a aplicação de técnicas de gerenciamento energeticamente eficientes e mediante o uso de recursos mais sustentáveis, os provedores de serviço de rede aumentariam sua competitividade. A competitividade pode aumentar tanto com redução de custos, quanto com uma possível atividade no mercado de crédito de carbono.

Por fim, visto que crescem as despesas e a pegada de carbono referentes ao setor de TIC, visto que a redução dessa pegada passa a gerar preocupações (GeSI, 2008) e visto que as redes correspondem a uma parte significativa das despesas desse setor, diversas soluções para tornar as redes energeticamente mais eficientes vêm sendo propostas (Gunaratne, 2005). As soluções variam desde projetos mais eficientes de componentes, até gerenciamento completo de sistemas de redes (Bolla, 2011a). Com as soluções, vem a necessidade de avaliar seus efeitos.

1.2 Objetivo

Este trabalho apresenta métodos para (i) a avaliação do impacto da adoção de técnicas verdes no tempo de vida de equipamentos e na disponibilidade da rede; e para (ii) a obtenção de um banco de experimentação que permita a realização de tal avaliação fazendo-se uso de emulação de soluções de gerenciamento verde.

Técnicas existentes relacionadas a sustentabilidade em sistemas de TIC exigem funcionalidades ainda não incorporadas na maioria dos equipamentos, sendo, portanto, essas próprias técnicas ainda não plenamente aplicáveis (Gupta, 2004; Gunaratne 2008; Chabarek e Barford, 2011; Bolla, 2011a). Quanto ao objetivo (i),

nota-se que essas técnicas afetam a rede de alguma forma além da intencionada, sendo o estudo de seus efeitos objeto deste trabalho.

Por exemplo, ver-se-á que se pode obter o consumo da rede, com boa precisão, por meio da utilização de modelos de consumo de energia dos dispositivos que compõem a rede. Entretanto, modelos assim ainda não são fornecidos diretamente pelos equipamentos. Ainda, os equipamentos não apresentam consumo linear com a carga, sendo seu consumo em muitos casos praticamente constante. Assim, avaliar futuras soluções verdes requer um ambiente de experimentação com algumas características específicas. Discutir e mostrar como um tal ambiente de experimentação pode ser montado corresponde ao objetivo (ii) desta dissertação. Nesse ambiente novas arquiteturas e soluções energeticamente eficientes de rede e novas funcionalidades verdes podem ser testadas.

Como efeitos colaterais da aplicação de técnicas verdes, será apresentado de que modo tais técnicas afetam o tempo de vida dos equipamentos ou a confiabilidade e disponibilidade da rede. O conhecimento dos efeitos sobre o tempo de vida é valioso, pois pode ser aplicado a decisões econômicas. A este respeito, no entanto, não será abordada logística reversa ou descarte, mas o impacto da operação no tempo de vida do equipamento. Os mecanismos que ocasionam falha nos dispositivos da rede serão destrinchados e será visto como avaliar os efeitos da nova operação, sustentável. Quanto ao outro efeito colateral, a confiabilidade afeta a experiência do usuário, ao passo que a disponibilidade do serviço prestado pela rede é usualmente expressa em acordos de nível de serviço. Avaliar como novas técnicas verdes as modificam também tem valor econômico agregado. A este respeito, será apresentado um novo método que, no cálculo da confiabilidade e disponibilidade da rede, inclui os efeitos de haver agora um dispositivo que ora dorme, ora acorda¹.

¹ Neste trabalho são utilizadas as expressões “estado dormiente” e “estar dormindo”, respectivamente, quando referindo-se ao *sleep mode* e ao estado de um dispositivo que nesse

A coordenação entre as funcionalidades energeticamente eficientes e seus efeitos colaterais pode ser mais bem efetivada com um sistema de gerenciamento de redes. Dadas as novas atribuições sustentáveis, este passa a receber o nome de sistema de gerenciamento energeticamente eficiente de redes. Aspectos da implementação desse sistema serão detalhados e será discorrido acerca dos desafios encontrados ao se implementar um ambiente de experimentação para sua validação, posteriormente discutindo os resultados. Tanto a arquitetura do sistema abarca tecnologia ainda não plenamente madura, quanto prevê o uso de equipamentos legados. Tal heterogeneidade deverá também ser contornada pelo ambiente de experimentação utilizado na avaliação.

1.3 Organização

O conteúdo deste trabalho é organizado da seguinte maneira: Este Capítulo 1 introduz o trabalho ao contextualizá-lo e apontar sua motivação, objetivos e organização. O Capítulo 2 apresenta o método aplicado às diferentes partes da pesquisa. O embasamento teórico e trabalhos relacionados são apresentados e discutidos no Capítulo 3. O Capítulo 4 discute impactos que a aplicação de técnicas verdes impinge ao tempo de vida dos equipamentos da rede, sugerindo um método para medi-los, ao passo que o Capítulo 5 apresenta como a aplicação de modo dormente afeta a confiabilidade e disponibilidade da rede. Uma arquitetura para gerenciamento orientado por política e atenta aos compromissos citados nos Capítulos 4 e 5 é apresentada no Capítulo 6. O Capítulo 7 apresenta uma implementação dessa arquitetura e exhibe como se obter um banco de testes para soluções equivalentes. O Capítulo 8 discute acerca dos resultados da avaliação do sistema aplicado ao ambiente de experimentação. Por fim, o Capítulo

modo se encontre. Ainda, aplicar-se-á o verbo “acordar” ou “despertar” quando se designando a transição de um dispositivo do modo dormente para o ativo.

9 conclui este trabalho e norteia trabalhos futuros na área de eficiência energética aplicada à TIC, no que tanga às redes de computadores.

2 Método

Neste capítulo descrevem-se aspectos relacionados ao método aplicado durante a pesquisa ora dissertada. A pesquisa se deu de 2 de janeiro de 2012 a dezembro de 2013, no LASSU (Laboratório de Sustentabilidade em Tecnologia da Informação e Comunicação), USP. As seções seguintes detalham mais pormenorizadamente a natureza metodológica dos ramos desta dissertação.

2.1 Do ambiente de experimentação para sistemas de gerenciamento energeticamente eficiente de redes

O ambiente de experimentação é resultado de uma pesquisa aplicada segundo o método hipotético dedutivo. Nesse método, “a solução é especificada e implementada por meio de um protótipo de modo a gerar resultados que possibilitem a validação da hipótese” (Miers, 2012). Além do ambiente de experimentação ser neste trabalho objeto desse método, pode ele mesmo servir de supedâneo na validação de protótipos de outras pesquisas aplicadas hipotético-dedutivas no âmbito de sustentabilidade em redes.

As referências elencadas durante o período de pesquisa bibliográfica não expressam como avaliar as técnicas, por exemplo de roteamento (mais discutidas no Capítulo 3), no âmbito da sustentabilidade, questão inicialmente abordada apenas em (Chabarek e Barford, 2011). Em (Januário, 2013) uma abordagem semelhante à de Chabarek e Barford foi ampliada para emulação de mais funcionalidades necessárias para redes sustentáveis. O raciocínio por detrás de como testar dada técnica de eficiência energética para redes de computadores foi sistematizado assim (Januário, 2013):

1. Definição do ambiente de experimentação: topologia da rede, perfil de tráfego, perfil de consumo dos roteadores, taxa de amostragem ou de tomada de aquisição de dados da rede.
2. Definição dos resultados da experimentação: ganhos de energia, sobrecusto na operação, mudança na disponibilidade e confiabilidade da rede, mudança no tempo de vida dos equipamentos.

2.2 Do método de avaliação de confiabilidade e disponibilidade

Será utilizado o trabalho de Amaral (2012) no desenvolvimento de uma nova técnica de se calcularem a confiabilidade e a disponibilidade de uma rede que aplique o modo dormente a alguns dispositivos. Aquele autor já demonstrou que há degradação gerada pelo modo dormente, propondo pioneiramente o método REASoN para avaliá-la. Esta dissertação mostra uma extensão ao método lá apresentado, motivada pelas definições de confiabilidade fornecidas por Johnson (1989), Callou (2012), Rosing (2007) e por Amaral (2012). Como resultado da extensão proposta, a demora para acordar, mesmo se algo menor, passa a ser mais impactante.

2.3 Do método de avaliação do tempo de vida dos dispositivos

O Capítulo 3 apresenta técnicas que pretendem levar sustentabilidade a redes de computadores por meio de eficiência energética na operação, prática que acarretaria redução de custos de operação. Mas eficiência energética não compreende o todo da sustentabilidade. Há de se considerar o ciclo de vida do equipamento. Assim, deve-se confirmar se há consistência interna na solução, se há harmonia entre a atuação focada apenas na operação e o todo da

sustentabilidade que envolve o equipamento. Para tanto, parte-se das seguintes premissas:

- Sustentabilidade engloba fabricação, operação e descarte dos equipamentos.
- As técnicas estudadas e descritas no Capítulo 3 atuam durante a operação

Dado isso, a pergunta de pesquisa a ser formulada é: Altera a atuação dessas técnicas, de algum outro modo, o todo do ciclo de vida dos equipamentos? Como medir?

O resultado da investigação é principalmente uma pesquisa bibliográfica. Sabe-se que a operação do equipamento se dá durante sua vida útil. Quanto menor a vida útil dos equipamentos, mais frequentemente é necessário repô-los e maior é o custo total de fabricação e descarte.

A definição sobre quais partes dos equipamentos de rede deveriam receber foco começou com a inspeção de termos de garantia. Por exemplo, o *data sheet* da Juniper (2011) declara que, para os sistemas Juniper *Networks* EX2200, EX3200, e EX4200, o *hardware* será essencialmente coberto pela garantia para sempre, mas nota “que a garantia da ventoinha e da fonte de alimentação são limitadas a 5 anos a partir da data de início”. Em (Juniper, 2013), declara-se que os roteadores lá abordados são providos com fonte de alimentação substituível no local de operação e com ventoinha quentemente trocáveis (*hot-swappable fan tray*), pelo bem de “funcionalidades de alta disponibilidade”. Com esses fatos em mente, pode-se deduzir que duas partes mecânicas, a ventoinha e a fonte de alimentação, são os componentes que mais falham nos dispositivos de interesse.

Sem embargo, isso não significa que partes eletrônicas devam ser deixadas de lado, como bem observado em (Srinivasan, 2004). Quer-se garantir que um novo padrão de operação no processador não o fará falhar com muito mais frequência, de modo que venha a falhar com frequência comparável às ventoinhas e fontes de alimentação. Por fim, a análise da fonte de alimentação será posta de lado, pois

não se lida aqui com a possibilidade de gerenciamento desse tipo de componente. Fica o estudo de tempo de vida então restrito ao âmbito de ventoinhas e processadores.

2.4 Considerações do capítulo

A avaliação de técnicas para sustentabilidade, mais especificamente para eficiência energética, traz desafios particulares. A cada desafio cabe uma abordagem diferente, conforme esteja seu estado-da-arte. Determinar um ambiente de experimentação é possível uma vez que se tenha entendido o que é comum entre técnicas de eficiência energética e o que estas exigirão dos equipamentos de rede. A tal tarefa é adequada pesquisa bibliográfica sobre essas técnicas e validação do ambiente de experimentação.

Uma vez que se tenha entendido o que as técnicas de eficiência energética ocasionam na rede, podem-se estudar seus efeitos. Efeitos em confiabilidade e disponibilidade, por já haverem sido apresentados no estado-da-arte, podem ser estendidos e revistos. Quanto aos efeitos no tempo de vida dos equipamentos, a estes cabem uma pesquisa bibliográfica e estudos iniciais de como beneficiar-se de seu conhecimento.

3 Trabalhos relacionados e embasamento teórico

Neste capítulo encontra-se embasamento para os demais capítulos, bem como uma apresentação de soluções para eficiência energética em redes de computadores. Assim, a Seção 3.1 descreve sistemas de roteamento sustentável, ao passo que a Seção 3.2 descreve tecnologias de *hardware* que, forçosamente, atuam localmente nos dispositivos. Ambas as seções descrevem técnicas de eficiência energética em redes de computadores. Entender os tipos de operação descritos nessas duas seções é importante na determinação tanto de como essas operações podem ser testadas, quanto de como elas podem afetar as redes. A Seção 3.3 apresenta técnicas de emulação que podem ser aplicadas às técnicas de eficiência energética, desta forma fornecendo embasamento teórico para o Capítulo 7. As Seções 3.4 e 3.5 fornecem embasamento teórico para o SustNMS (Capítulo 6), um sistema de gerenciamento orientado por política e voltado a eficiência energética. Esse sistema é avaliado no estudo de caso do Capítulo 8. As Seções 3.6 – 3.9 fornecem embasamento teórico sobre confiabilidade, cálculo tempo médio até a falha (MTTF) e assuntos correlatos, conceitos importantes para os Capítulos 4 e 5. Considerações finais do conteúdo deste capítulo são apresentadas na Seção 3.10.

3.1 Sistemas de gerenciamento voltados a eficiência energética

Abordagens como o Sistema de Gerenciamento de Energia (*Energy Management System*) (Maciá-Pérez, 2009) propõem sistemas de gerenciamento de consumo de energia que requerem informação sobre a rede toda. Os autores sugerem gerenciar a energização e desenergização de elementos de rede por meio da inclusão de serviços nos equipamentos, como *Web Services* específicos para

esse fim. A solução proposta desabilita enlaces ociosos e, se possível, mesmo roteadores ou comutadores.

O trabalho de Chaudhari (2012) propõe um sistema de gerenciamento de redes energeticamente eficiente que utiliza uma MIB com suporte a mecanismos de eficiência energética. Decisões sobre que componentes desligar são baseadas na topologia da rede e em políticas especificadas pelo usuário. Os autores propõem a adição de um novo estado à MIB definida no RFC 2836, a saber, o modo dormente. Tal solução pode ser implementada quando todos os equipamentos de rede tiverem suporte ao novo estado e o protocolo SNMP (*Simple Network Management Protocol*) ou equivalente habilitado, não sendo possível atualmente pô-la em prática de modo completo.

Um mecanismo centralizado de engenharia de tráfego com foco em eficiência energética foi descrito em (Zhang, 2010), o GreenTE (acrônimo de *Green Traffic Engineering*). Esse mecanismo visa a calcular rotas usando a topologia da rede e sua matriz de tráfego. O objetivo é maximizar a eficiência energética ao colocarem-se em modo dormente placas de linha (*linecards*) dos dispositivos da rede e, ao mesmo tempo, manter-se o desempenho da rede a níveis desejados. A abordagem adotada consiste na formulação de um problema de Programação Inteira Mista² (Schrijver, 1998) no qual o objetivo a ser maximizado é o ganho de energia. Concomitantemente, requisitos de desempenho como a utilização máxima de enlace e o atraso da rede são considerados restrições à solução. Os autores observam que a formulação do problema traz similaridades à engenharia de tráfego tradicional, contrapondo objetivos distintos das duas formulações. Ao passo que a primeira tenta concentrar o tráfego a um subconjunto dos enlaces, a segunda, tradicional, tenta espalhá-lo uniformemente por todos os enlaces. O algoritmo do GreenTE objetiva agregar as portas livres do equipamento o máximo possível nas mesmas placas de linha, de sorte que não apenas cada porta livre

² Programação Inteira Mista é uma especificidade da Programação Linear (PL) em que algumas variáveis são inteiras e outras, contínuas. PL, por sua vez, é utilizada em problemas onde há restrições e objetivos, ambos com função linear.

possa ser desligada, mas também toda a estrutura da placa de linha, economizando mais energia.

O projeto MiDORi (*Multi-layer, path, and resources Dynamically Optimized Routing*) (Takeshita, 2011) propõe um mecanismo para computação de caminhos que calcula topologias ótimas que acomodem todo o tráfego. Cada roteador ou comutador monitora o tráfego de cada caminho MPLS³, então o mecanismo de cálculo de rotas entra em ação. O projeto efetua a sinalização de desligamento ou ativação dos enlaces através de uma extensão própria ao MPLS.

Devido à abrangência dos respectivos projetos, o Green OSPF e o ElasticTree merecem subseções próprias, sendo descritos abaixo.

3.1.1 Green OSPF

Cianfrani (2010) propõe uma solução de roteamento que visa a economizar energia na rede. Ele busca uma solução que seja compatível com protocolos clássicos de roteamento, como o OSPF (*Open Shortest Path First*). A solução baseia-se no OSPF, de modo a ser facilmente implementada e obter informação sobre a topologia da rede. Ao propor um roteamento voltado ao consumo energético da rede, Cianfrani contribui para o estado-da-arte avançar em direção a um OSPF verde aplicável às atuais redes IP.

O protocolo OSPF demanda que cada roteador calcule sua árvore de rotas mais curta através do algoritmo de Dijkstra (Skiena e Revilla, 2003). A ideia do algoritmo de Cianfrani, o *Energy-Aware Routing* (Roteamento Atento a Energia), é que apenas um subconjunto de roteadores sejam usados para definir as rotas, de modo que, como um todo, seja reduzida a quantidade de enlaces utilizados no roteamento do tráfego.

³ A Seção 3.5 conceitua o MPLS.

Os trabalhos sobre Green OSPF não tratam casos de distúrbios de tráfego que venham a ocorrer enquanto o sistema de gerenciamento recalcula caminhos. De fato, essa falta é esperada, pois esse é um problema com qualquer protocolo de roteamento, não sendo o OSPF uma exceção⁴.

3.1.2 ElasticTree

O sistema ElasticTree (Heller, 2010) aplica inativação dinâmica de equipamentos de rede, focando em economia de energia elétrica. Esse sistema dá um passo em direção à desejada proporção entre consumo de energia e carga de trabalho da rede. Isso não é obtido em nível individual de cada equipamento, mas de sua operação dentro do contexto da rede onde o equipamento se insere. Assim, ocorre o desligamento total de equipamentos e não de suas placas de rede ou interfaces de linha. Tal desligamento ocorre de sorte que a rede como um todo resulte energeticamente mais proporcional em relação a sua carga de trabalho.

Na busca para manter a rede ainda capaz de lidar tanto com a carga oferecida, quanto com a nova carga do futuro imediato, o ElasticTree lança mão de parâmetros de margem de segurança. Tais parâmetros controlam compromissos entre desempenho, robustez e energia. Heller também busca analisar diferentes métodos de roteamento, operando com o ElasticTree em ambientes de *data centers*. Sua abordagem é baseada em modelos de consumo de energia elétrica e implementada em OpenFlow. Seus modelos são baseados em medidas obtidas em sistemas de *hardware* reais, incluindo medidas referentes ao tempo necessário para o dispositivo e seus enlaces serem ligados ou desligados. Como exemplos, citam-se o caso do comutador Quanta, para o qual obteve-se o menor tempo para acordar de cerca de 30 segundos; e o caso do comutador HP, cujo valor chegou a

⁴ Devo as observações desse parágrafo a Catalin Meirosu.

3 minutos. De modo geral, energizar e desenergizar individualmente as interfaces demora de 1 a 3 segundos.

A qualidade do resultado obtido pelo ElasticTree depende do tempo gasto no levantamento estatístico da rede, no tempo de cálculo da solução de roteamento, e no tempo gasto para acordar um dispositivo e habilitar seus enlaces. Quanto maior a soma desses tempos, pior é a qualidade da resposta do ElasticTree à ocorrência de picos de tráfego.

3.2 Funcionalidades de equipamentos voltadas a eficiência energética

As técnicas para redução de consumo energético que atuam localmente nos equipamentos de redes foram categorizadas em três grupos por Bolla (2011a): reengenharia, adaptação dinâmica e aplicação de modo dormente.

A primeira categoria refere-se ao projeto de equipamentos que sejam internamente mais eficientes, o que comumente requer a aplicação de materiais e componentes também eficientes. Tecnologias puramente ópticas são a vedete desta categoria, que também conta com novas tecnologias de memória e de processamento de pacotes. Bolla (2011a) também afirma que componentes puramente ópticos poderiam fornecer banda com largura de *terabits* e manter o desperdício de energia bem menor que o dos dispositivos atuais.

A segunda categoria de técnicas verdes refere-se à adaptação dos equipamentos de rede segundo o tráfego corrente e requisitos de serviço. Essas técnicas requerem interação com o *hardware* do equipamento, sendo as duas principais o escalamento de energia e a lógica ociosa⁵. Essas técnicas ainda não são comuns em equipamentos de rede. Com o escalamento de energia pode-se mudar a frequência de operação do processador e das interfaces de equipamentos em

⁵ *Power scaling* e *idle logic*, no inglês.

função da demanda de tráfego, a fim de se otimizar o consumo de energia. Gupta (2003) e Nedevschi (2008) modelam o comportamento desse tipo de técnica, conforme discutido mais adiante na Seção 3.2.2. Tipicamente, tal efeito é obtido pelo gerenciamento da frequência do *clock* ou da voltagem do equipamento. A lógica ociosa refere-se ao desligamento, quando possível, de subsistemas do equipamento. (Por exemplo, desligamento do subsistema responsável pelo plano de dados do dispositivo que apresente tráfego nulo.)

A terceira categoria de técnicas verdes refere-se ao uso do modo dormiente. Esta categoria exige tanto que os equipamentos de rede operem em algum modo de baixo consumo de energia, quanto que possam, mesmo estando num tal estado, trocar e interpretar mensagens de comando, como requisições para acordar. Um dispositivo dormiente perde sua função na rede e, segundo Bolla (2011a), demora um tempo não negligenciável para acordar. As implicações dessa demora e um método de aferi-las serão estudados no Capítulo 5. Há também abordagens que buscam novas formas de desligar apenas subsistemas dos dispositivos, como placas de linha (Allman, 2007), (Chabarek, 2008), (Gupta, 2003) e (Heller, 2010).

Nos quatro parágrafos acima se discutiram técnicas que agem localmente, em um único equipamento. Há também abordagens que visam a, como um todo, gerenciar a rede de modo a torná-la mais energeticamente eficiente. Tais técnicas atuam no sentido de forçar o consumo da rede a variar mais linearmente de acordo com a carga total da rede, ainda que, individualmente, os dispositivos não apresentem tal comportamento linear.

Entretanto, as abordagens de gerenciamento energeticamente eficientes existentes (Maciá-Pérez, 2009), (Chaudhari, 2012), (Zhang, 2010), (Takeshita, 2011) e a extensão de Chabarek e Barford (2011) não propiciam um ponto de vista de gerenciamento guiado por políticas, e as soluções correlatas não incluem o suporte à aplicação de decisões baseadas em política. Isso está em desacordo com a tendência de gerenciamento orientado por política (Oltsi e Jon, 2012) e seus benefícios, apontados em (Strassner, 2002; e Strassner 2003). Ainda, esta

dissertação levanta compromissos acarretados por operações voltadas a sustentabilidade, os quais apenas agravam a necessidade de gerenciamento orientado por política.

3.2.1 ACPI

A ACPI (*Advanced Configuration and Power Interface*, Configuração Avançada e Interface de Potência) especifica os estados energéticos de processadores e define uma interface de comunicação que permite determinar ou descobrir quais são os estados energéticos correntes desses processadores (ACPI, 2010). Foi especificada por um consórcio de empresas para facilitar a interoperabilidade no gerenciamento de energia efetuado diretamente por sistemas operacionais. Para cada diferente estado energético, a ACPI determina quais subsistemas do processador devem estar ligados e, desta forma, qual a potência dissipada. (A potência é determinada apenas indiretamente, pois depende do fabricante do equipamento).

Os tempos de transição entre os estados energéticos da ACPI, em particular para ir ou voltar do modo dormiente, dependem da implementação de cada fabricante. Para fins de emulação ou simulação, é importante que se tenha informação sobre tais tempos de transição. Por exemplo, o tempo de transição para despertar de um processador Intel da geração Haswell varia de 1 a 3 milissegundos⁶.

Num esforço semelhante, a *Internet Engineering Task Force* (IETF, Força-Tarefa de Engenharia na *Internet*) tem criado padronizações para os dados relacionados ao consumo energético de equipamentos de rede. O documento de esboço referente à Base de Informação para Monitoração de Potência e Energia (*Power and Energy Monitoring MIB*) especifica um modelo para suporte a gerenciamento

⁶ Observações realizadas por Catalin Meirosu, *Ericsson Research*, Suécia.

de energia (Chandramouli, 2011). Tal modelo prevê a existência, entre os estados ativo e inativo de um equipamento, de vários modos intermediários para redução do consumo de energia.

Com a proliferação de ACPI nos processadores e os esforços de padronização da IETF, não é exagero esperar que dispositivos de rede logo venham a ter mecanismos semelhantes (Gupta, 2004). Nesse cenário, um método de se avaliar o tempo de vida dos dispositivos e mudanças na disponibilidade da rede ganha importância.

3.2.2 Taxa Adaptativa de Enlace

A taxa adaptativa de enlace (ALR, *Adaptive Link Rate*) é um meio de reduzir o consumo de energia elétrica por equipamentos de rede Ethernet. A redução é obtida pela mudança da taxa dos enlaces em consonância com a utilização requerida (Gunaratne, 2005). Essa técnica foi criada com o intuito de utilizar taxas de enlace atualmente existentes (10Mbps, 100Mbps, 1Gbps e 10Gbps). Ela se baseia no fato dos enlaces Ethernet serem, em média, subutilizados.

O funcionamento do ALR é regido por dois fatores. O primeiro é o mecanismo que funciona nos dispositivos dos extremos do enlace e determina como a taxa de transmissão é alterada. O principal exemplo de mecanismo é o *Rapid PHY Selection* (Seleção Rápida na Camada Física), estudado em (Blanquicet e Christensen, 2007), baseado em mensagens de camada de enlace. O segundo fator é a política que determina quando a taxa deve ser alterada. É importante que a política maximize o tempo que o enlace opera em taxa baixa e minimize o atraso de pacotes ocasionado.

Gunaratne (2008) estudou analiticamente o compromisso entre ganhos de energia e atraso de pacote envolvidos na operação do ALR. Observou que, com uma

utilização média do enlace de 5% ou menos, é possível operar a 100Mbps em mais de 80% do tempo, e ainda gerar um atraso menor que 0,5 milissegundos.

3.2.3 Coalescimento sincronizado no comutador

A técnica de coalescimento no comutador (SSC, *Synchronized Switch Coalescing*) (Mostowfi e Christensen, 2011) é direcionada para redes Ethernet locais, de escritórios pequenos e casas (as chamadas rede SOHO, ou seja, *Small or Home Office*), pois nessas redes encontram-se os dispositivos mais ociosos e pouco utilizados. O SSC foi desenvolvido para aproveitar a usual baixa carga e os incontáveis períodos ociosos que os dispositivos de rede SOHO encontram. Ao mudar o comportamento desses dispositivos de completamente ligados para modos de energia reduzida, Mostowfi e Christensen pretendem diminuir o consumo de energia elétrica desse tipo de rede, atualmente de cerca de 8 TWh por ano, apenas nos Estados Unidos da América.

A técnica SSC consiste em impedir que os enlaces conectados a dado comutador encaminhem tráfego durante certo tempo. Em seguida, as interfaces dos dois extremos de cada enlace entram em modo de baixo consumo energético, de sorte que vários componentes do comutador em questão possam entrar em modo dormente. O resultado é uma minimização do consumo de energia do comutador, enquanto ainda se mantendo o desempenho da rede. A ideia do coalescimento sincronizado no comutador remonta à ideia da técnica denominada Ciclo de Potência Pausado (*Pause Power Cycle*), previamente proposta e prototipada por Blanquicet (2008). Em seus testes, Mostowfi e Christensen (2011) utilizaram os tempos de transição entre os estados C0 e C6 da ACPI de um processador Intel i5 multinuclear, cerca de 2 milissegundos.

3.3 Ambientes de experimentação

Um ambiente de experimentação no qual as soluções supracitadas possam ser validadas e comparadas é tão importante quanto as próprias soluções. Ambientes para pesquisa experimental, como o simulador ns-2 (*The Network Simulator*, 2010) e o Emulab (Emulab, 2010), são bem conhecidos pela comunidade de pesquisa.

Extensões para ambiente de experimentação e que aprimoram roteadores baseados em GNU/Linux por meio de configurações eficientes foram propostas por Chabarek e Barford (2011). As extensões permitem diversos experimentos em roteamento verde. São elas: (i) um monitor que reporta em tempo real uma estimativa, validada por Chabarek e Barford como altamente acurada, do consumo de energia elétrica de um dispositivo, e (ii) um modelador de tráfego que, através do controle do fluxo de saída, pode emular alguma funcionalidade que venha a mudar a taxa dos enlaces do dispositivo e, com isso, alterar o consumo de energia de tal dispositivo. O trabalho também relata a dificuldade de se obterem medidas de energia em tempo real nos equipamentos existentes, problema para cuja solução os autores introduziram perfis de energia nos roteadores. Um perfil de energia relaciona a potência gasta por um dispositivo quando este lida com determinada carga. Chabarek e Barford (2011) confirmaram que abordagens mais granulares acarretam maior ganho de energia quando o custo de energia de cada transição entre estados é baixo e o tráfego varia, desta forma apontando a importância da precisão nos modelos de consumo. Esses perfis são praticamente invariáveis, refletindo os resultados encontrados em medições. Nota-se que tal invariância está em desacordo com as capacidades previstas e desejadas para os roteadores futuros, dissonância a qual o presente trabalho busca dirimir.

Para validar suas extensões e demonstrar sua utilidade, Chabarek e Barford (2011) montaram um ambiente de experimentação com comutadores comuns e baseados em Linux. Em um de seus experimentos, eles utilizaram uma das

extensões para aumentar a largura de banda do comutador à medida que o tráfego demandasse. As extensões por eles desenhadas ainda atuam no eixo de desempenho e do consumo energético, desta forma não considerando compromissos, recentemente levantados, que possam existir com o tempo de vida do dispositivo e com a confiabilidade e disponibilidade da rede.

3.4 Gerenciamento orientado por política

No contexto de redes voltadas a eficiência energética, sistemas de gerenciamento orientado por políticas ganham mais espaço. Como apontado anteriormente, algumas técnicas verdes colocam dispositivos da rede em modo dormente, de sorte que ele perca sua presença de rede. Com a perda de presença de um ou mais dispositivos, pode-se ter a qualidade do serviço prestado comprometida. Ainda, como também se verá neste trabalho, ligar e deligar um dispositivo mais frequentemente acarreta mudanças na confiabilidade do próprio dispositivo e, destarte, da rede também. A partir de um sistema de gerenciamento orientado por política, o compromisso, inerente às técnicas verdes, entre ganho de energia e perda de qualidade de serviço pode mais facilmente ser gerenciado.

O gerenciamento orientado por política, quando esta é descrita em nível de abstração mais alto, sem envolvimento de comandos específicos de equipamentos, fornece um controle mais acessível a gerentes de redes ou de sistemas de TIC (Tecnologia da Informação e Comunicação) onde tais redes estão inseridas. Uma rede deve atender aos níveis de serviços contratados pelos clientes, definidos nos acordos de nível de serviço. Em uma rede constituída por equipamentos de tipos e fabricantes distintos, saber como configurar cada tipo de equipamento pode ser um empecilho, de modo que, se o controle puder ser definido em nível de abstração mais alto, a tarefa é facilitada. Quanto às abordagens citadas, argumenta-se aqui que com elas não é possível uma avaliação abrangente e em tempo real sobre os impactos que práticas verdes,

como transitar entre diferentes estados de energia, acarretam à confiabilidade da rede e dos dispositivos e, de modo geral, à qualidade de serviço. Embora soluções existentes (Chabarek e Barford, 2011), (Heller, 2010) e (Broch, 1998) propiciem suporte de ambiente de experimentação para que se avaliem funcionalidades voltadas a eficiência energética, não se encontrou na literatura especializada algum sistema de gerenciamento energeticamente eficiente com abordagem integrada e baseada nesse tipo de ambiente.

Para lidar com tais questões, necessitam-se um ambiente de experimentação de sustentabilidade holístico e um sistema de gerenciamento de redes energeticamente eficiente, preferivelmente de decisões centralizadas. Centralidade na arquitetura sustentável é útil quando não se dispõe localmente de toda informação sobre confiabilidade e disponibilidade, consumo de energia e tráfego. Tal sistema de gerenciamento de redes requer conhecimento constante e atual do estado da rede para que ele possa avaliar os compromissos inerentes a quaisquer decisões, como sobre qual estado ACPI acarretaria maior confiabilidade enquanto ainda economizando 5% de energia. O sistema também deve coordenar funcionalidades energeticamente eficientes numa rede heterogênea. Um sistema de gerenciamento de redes consegue perfazer essa tarefa beneficiando-se de perfis de consumo de energia, os quais são modelos que relacionam o consumo energético de um dispositivo ao tráfego demandado. Nota-se que uma rede heterogênea provavelmente contém equipamentos legados, aos quais pode faltar suporte a funcionalidades energeticamente eficientes. Não obstante, pode-se modelar tal suporte por meio de modelos de consumo de energia. Ainda, pode-se prover um ponto de vista de nível mais alto de gerenciamento uma vez que o sistema proveja suporte a políticas e permita análise do compromisso entre ganhos de energia e a degradação sentida no nível de serviço.

Costa et al. (2012a) apresentou uma arquitetura que satisfaz aos requisitos discutidos, ou seja, uma arquitetura holística e baseada em política. Com base nessa arquitetura, um protótipo de um sistema para gerenciamento energeticamente eficiente de redes e orientado por política, o SustNMS, foi

desenvolvido. Aspectos de sua implementação e avaliação são mostrados em (Januário, 2013). O sistema considera garantia de qualidade de serviço e assegura resposta rápida a falha e a aumento repentino de tráfego. Para cada decisão sobre em que estado um nó deve se encontrar (por exemplo, se em modo normal ou de consumo reduzido), deve ser efetuada uma análise em tempo real sobre a confiabilidade da rede, de modo que se preveja mais clarificação sobre os potenciais impactos de se aplicar a decisão, conforme apresentado em (Amaral, 2012). A arquitetura do SustNMS será discutida mais pormenorizadamente no Capítulo 6.

A arquitetura de Costa (2012a) prescreve o uso de política no gerenciamento e contempla a possibilidade dessas políticas utilizarem parâmetros de confiabilidade, disponibilidade, desempenho e ganhos energéticos. Entretanto, tal arquitetura não prescreve como tais políticas devem ser expressas. Para esta dissertação, as políticas utilizadas foram embutidas no código do SustNMS (Capítulo 6).

3.4.1 Aspectos do gerenciamento orientado por política

O gerenciamento de redes orientado por política encerra um paradigma de gerenciamento que separa operações administrativas de outras operações básicas da rede (Rana, 2011). Esse paradigma ganhou importância nas últimas duas décadas, mostrando que a dicotomia entre operações administrativas e operações básicas é um mecanismo flexível e robusto (Carvalho, 2012). Boros (2000) identificou especial interesse do gerenciamento de redes orientado por política nas seguintes áreas de operação de redes:

1. Políticas de gerenciamento de desempenho;
2. Políticas de controle de acesso e segurança;
3. Políticas de qualidade de serviço;

4. Políticas de administração e configuração;
5. Políticas de gerenciamento de falhas.

Quando usado com um contínuo de política, as regras administrativas, de nível mais alto, requerem maior interpretação pelo sistema de gerenciamento, pois são mais gerais e abstratas. As regras de nível mais baixo são mais específicas, prescrevendo que ações os dispositivos devem tomar como resposta a alguma regra administrativa, ou de que modo eles devem ser configurados.

Independentemente se utilizado com um contínuo de política ou não, a arquitetura dum sistema orientado por política deve gozar de (IETF, 1999) :

1. Habilidade para definir e atualizar regras;
2. Habilidade para armazenar e recuperar regras;
3. Habilidade para entender os critérios condicionais duma regra;
4. Habilidade de executar as ações determinadas por uma regra quando os critérios condicionais forem satisfeitos.

3.4.2 Contínuo de Política

Quanto à tradução de políticas, Rana (2011) adverte sobre a complexidade da transformação de conceitos relacionados a políticas abstratas para políticas concretas. Ainda, atenta que, apesar das linguagens de política à época não proverem os necessários modelos semânticos, a existência destes poderia refrear o processo de tradução. Rana (2011) refere-se a tais modelos semânticos tendo em mente o contínuo de política apresentado em (Davy, 2007), para o qual nota que particularmente difícil é a tradução de políticas declarativas a políticas executáveis.

Políticas declarativas de alto-nível, especialmente para sistemas de rede, são por natureza abstratas, ao passo que linguagens executáveis, ou seja, de baixo-nível,

são comumente específicas para o domínio em que se inserem, contendo sintaxe formal atrelada e sendo de aplicação limitada (Rana, 2011). Essa especificidade também é apontada por Beller (2004). Por fim, observa-se que Rana (2011) simplifica a tradução de políticas. Para tanto, abre mão da completude dos níveis do contínuo de política, fazendo uso apenas dos níveis de usuário e de dispositivo.

Rubio-Loyola (2006) nota que o refinamento de política é uma área chave e não resolvida do gerenciamento de redes orientado por política. O método de refinamento que ele estuda é voltado a refinamento de objetivos. Ou seja, ele parte daquilo que se deseja realizar e define o que precisa ser feito, em outro nível de abstração, para que aquilo de fato se realize.

3.5 O protocolo MPLS

O MPLS (*Multiprotocol Label Switching*) torna transparente para os equipamentos a infraestrutura da rede. Assim, infraestruturas de ATM (*Asynchronous Transfer Mode*), *Frame Relay*, Voz e IP podem ser utilizadas de forma mais integrada. Essa característica das redes MPLS é conhecida como *AToM (Any Transport over MPLS, Qualquer Transporte sobre MPLS)*.

Além do fato de comutar pacotes MPLS ser uma operação potencialmente mais rápida que comutar pacotes IP, há outras vantagens no emprego de MPLS numa rede. Dentre os benefícios citados em (Ghein, 2006), destaca-se a possibilidade de implementação de (i) uma rede em cujo interior não funcione o BPG (*Border Gateway Protocol*), (ii) um modelo ponto-a-ponto para MPLS VPN (*Virtual Private Newtok, Rede Virtual Privada*) e (iii) mecanismos de engenharia de tráfego.

Por meio de (i) consegue-se aliviar a carga a ser processada pelos roteadores que não sejam de borda. Lidar com essa carga pode ser tão custoso quanto lidar com mais de 150.000 rotas (Ghein, 2006). Deste modo, o benefício (i) também pode aliviar os custos dos equipamentos de núcleo, fazendo-os requisitar menos

memória. A possibilidade (ii) dá margem a uma vasta gama de métodos de refinamento e de gerenciamento de qualidade de serviço. Primeiro, associa-se cada usuário a algum dos níveis de serviço disponíveis, depois explora-se a qualidade da VPN oferecida aos usuários.

A partir da engenharia de tráfego por MPLS (iii) pode-se dar um passo adiante e implementar Rerroteamento Rápido (FRR, *Fast ReRouting*). Explorar-se-á FRR no ambiente de experimentação descrito no Capítulo 7. A técnica de Rerroteamento Rápido ocorre em menos que 50 milissegundos, duração que Ghein (2006) observa ser rápida até mesmo para os padrões do ano em que publicou seu trabalho.

A subcamada do protocolo MPLS, ou seja, sua pilha de rótulos, situa-se no pacote entre as cargas das camadas 3 e 2 do modelo OSI (*Open System Interconnection*), o que rende ao MPLS também a alcunha de “protocolo de camada 2.5” (leia-se “dois e meio”). A um grupo ou fluxo de pacotes encaminhados pelo mesmo caminho e de modo idêntico chama-se Classe de Equivalência de Encaminhamento (FEC, *Forwarding Equivalence Class*). Os pacotes que se enquadram na mesma FEC compartilham também o mesmo rótulo MPLS. Note-se que valores distintos em certo campo do cabeçalho MPLS (o campo EXP, que indica experimentação) podem acarretar diferentes tratamentos quanto ao encaminhamento, de modo que, se por um lado pacotes de mesma FEC tenham rótulo igual, por outro lado rótulos iguais não necessariamente acarretam mesma FEC.

O caso mais simples de implementação de uma rede MPLS é o de IPv4 sobre MPLS. O ambiente de experimentação apresentado no Capítulo 7 faz uso dessa montagem.

Os rótulos MPLS não precisam ter uma significância global na rede. Sem embargo, eles são relevantes na comunicação entre dois roteadores adjacentes. Podem-se-os distribuir de dois modos:

1 – como apêndices em um protocolo de roteamento IP existente, o que requer que cada protocolo seja adaptado; ou

2 – com seu próprio protocolo, sendo este usualmente *Tag Distribution Protocol*, *Label Distribution Protocol*, ou *Resource Reservation Protocol*.

Os roteadores com suporte a MPLS são conhecidos como LSR (*Label Switch Router*, Roteador de Troca de Rótulo). Estes podem ser de ingresso, de egresso ou intermediários. Roteadores de ingresso colocam o cabeçalho MPLS, usualmente já com rótulo inicial, nos pacotes. Roteadores de egresso livram o pacote de qualquer carga MPLS, ou seja, retiram-lhe o cabeçalho MPLS. Roteadores intermediários podem adicionar, tirar ou trocar rótulos.

O MPLS é a solução adotada para cumprimento de um dos requisitos do ambiente de experimentação para soluções sustentáveis apresentado no Capítulo 7. Esse protocolo foi adotado por permitir que alguma rota predeterminada seja aplicada a certo fluxo de pacotes.

3.6 Conceito de tempo de vida e a distribuição de Weibull

Com respeito a *hardware*, uma boa aproximação para a evolução do tempo médio até a falha de um equipamento é a curva da banheira, exemplificada na Figura 1. A curva é composta por três períodos distintos da operação do equipamento. Os três períodos são: a) período de falhas precoces, quando os dispositivos falham devido a defeitos de fabricação, b) período da vida útil do equipamento, quando as falhas ocorrem aleatoriamente e com taxa constante, e c) período de falhas de desgaste ou degenerescência (Klutke, 2003). O desgaste em equipamentos eletrônicos é ocasionado por fadiga ou depleção de materiais (por exemplo, elétrons), podendo ocorrer indesejadamente durante a operação do equipamento. Quando ocorre uma falha, os fabricantes trocam o produto vendido ou alguma parte dele, de modo a estenderem sua vida útil.

Figura 1- Curva da banheira, mostrando a evolução da taxa de falha de um dispositivo

Com respeito a falha, para descrever cada período da vida do equipamento é apropriado o uso de uma função com distribuição de Weibull. Essa distribuição utiliza um fator de forma, de modo que, variando-o, podem-se modelar as três diferentes distribuições da curva da banheira (Neto e Cymbalista, 2006). Ainda, para fenômenos de desgaste, Weibull é um modelo mais acurado que a distribuição exponencial (Sidharth e Sundaram, 2004), justamente por permitir uma taxa de falhas não constante. A Equação 3.1 descreve a função de distribuição ou repartição para uma distribuição de Weibull,

$$F(t) = 1 - e^{-(t/\alpha)^\beta} \quad (\text{Equação 3.1})$$

Onde $F(t)$ representa a probabilidade de uma falha acontecer até o tempo t e α é o parâmetro de escala para o percentil de 63,2 (ou seja, $t = \alpha \leftrightarrow F(t) = 0,632$). Alterar o valor do parâmetro de forma β provê diferentes formas para a distribuição, daí a flexibilidade do modelo de Weibull. Por exemplo, se $\beta = 1$, $F(t)$ coincide com a função de repartição de uma distribuição exponencial, a qual é relacionada ao período central da curva da banheira, que apresenta taxa de falha constante. Quando $\beta < 1$, $F(t)$ está relacionada ao início da curva da banheira, quando há falhas precoces. Ao longo desse período a taxa de falhas decresce. Ao contrário, para $\beta > 1$, $F(t)$ se relaciona ao final da curva da banheira, quando se percebem as falhas por degenerescência. A distribuição de Weibull é bastante

utilizada para avaliar o tempo de vida de equipamentos de TIC (Bussab, 2008). Na distribuição de Weibull, $R(t) = e^{-(t/\alpha)^\beta}$ e $Z(t) = (\beta/\alpha) \cdot (t/\alpha)^{\beta-1}$

3.7 Modelo de Márkov

O modelo de Márkov de um sistema codifica em si as probabilidades de transição entre os diferentes estados lógicos do sistema (Johnson, 1989). Os diferentes estados possíveis do sistema são elencados a partir de características relevantes do próprio sistema. Exemplos dessas características são a informação sobre o sistema estar ou não falho, ou a informação sobre quais submódulos do sistema estão funcionando e quais não estão.

O modelo usual de Márkov para cálculo de confiabilidade e disponibilidade utiliza taxas de falha ou reparo que sigam uma distribuição exponencial. As taxas são codificadas nas transições entre os estados. A partir do modelo de Márkov de confiabilidade ou do de disponibilidade definem-se equações diferenciais sobre a probabilidade do sistema estar em cada estado (Trivedi e Selvamuthu, 2008). Resolvidas as equações, pode-se calcular, por exemplo, a confiabilidade do sistema, o que corresponde ao cálculo do tempo esperado para que ele atinja algum estado de falha.

A partir das probabilidades de transição (no caso, relacionadas às taxas de falha e de reparo), dado o estado em que o sistema se encontre pode-se prever em que estado o sistema estará no futuro. Nessa predição é irrelevante a informação sobre estados anteriores em que o sistema tenha passado até que chegasse ao estado atual, a partir do qual a predição é feita. Por gozar dessa propriedade o modelo de Márkov é dito sem memória.

No contexto de redes, uma utilidade do modelo de Márkov é representar os estados lógicos do sistema quanto aos equipamentos de rede (roteadores e

comutadores) ou quanto aos enlaces. No primeiro caso, incluem-se no modelo apenas taxas de falha e reparo dos, por exemplo, roteadores, e cada estado denota quais roteadores estão funcionando, quais estão falhos e, neste trabalho, quais estão dormindo. No segundo caso, incluem-se no modelo apenas taxas de falha e reparo dos enlaces, e cada estado denota quais enlaces estão funcionando, etc. Uma terceira possibilidade seria modelar tanto os enlaces quanto os equipamentos.

Neste trabalho, apresentam-se modelos de Márkov relativos (i) a falha e dormência apenas de roteadores e (ii) a falha e dormência apenas de enlaces. Apresenta-se também como tais modelos podem ser aplicados no cálculo de confiabilidade e disponibilidade de redes voltadas a eficiência energética.

3.8 Confiabilidade, disponibilidade e taxa de falhas

A confiabilidade de um sistema é dita por Johnson (1989) ser “uma função do tempo, definida como a probabilidade condicional de que o sistema operará corretamente por todo o intervalo $[t_0, t_1]$, posto que o sistema operava corretamente no tempo t_0 ”. Para Callou (2012), “confiabilidade é a probabilidade de que o sistema entregará um conjunto de serviços durante dado período de tempo”, sendo o sistema dito tolerante a falhas “se ele não falha mesmo quando há componentes faltosos”. Por fim, Rosing (2007) declara que “a confiabilidade de um sistema é a função $R(t)$ de probabilidade, definida no intervalo $[0, \infty]$, de que o sistema operará corretamente, sem reparo, até o tempo t ”. As três definições acima são cônsonas, podendo ser resumidas, nas palavras de Amaral (2012), assim: “confiabilidade é a probabilidade do sistema operar corretamente durante todo um intervalo de tempo”. Ainda, do mesmo modo que se pode deduzir pelas definições de Callou, Amaral declara que não necessariamente um sistema confiável é tolerante a falhas.

A disponibilidade de um sistema é dita por Johnson (1989) ser "uma função do tempo, definida como a probabilidade do sistema estar operando corretamente e estar capaz de perfazer suas funções no instante de tempo t ". Para Callou (2012), "disponibilidade [...] quantifica o misto efeito de falha e reparo em um sistema". Nota-se que este conceito de reparo não fora utilizado por Callou, nem por nenhum outro autor, na definição de confiabilidade. Assim, espera-se que seja alcançável para a expressão da confiabilidade uma fórmula sem um termo correspondente a reparo. De fato, Rosing excluiu qualquer reparo de sua formulação de confiabilidade.

Confiabilidade quantifica a expectativa de um sistema operar corretamente *até* certo momento. Disponibilidade quantifica a expectativa de um sistema estar operando corretamente *em* dado momento. Assim, um sistema pode ter disponibilidade alta mesmo quando experimenta constantes períodos de falha, tendo, portando, baixa confiabilidade, desde que os períodos de falha sejam relativamente curtos. Um sistema com períodos de falha curtos tem reparos que ocorrem rapidamente. Shooman (2001) observa que, quanto mais indetectável o período de inacessibilidade, maior a disponibilidade.

A confiabilidade de um sistema para o instante t_0 pode ser matematicamente expressa como em (Neto e Cymbalista, 2006)

$$R(t_0) = P(T > t_0) = 1 - F(t_0) = \int_{t_0}^{\infty} f(t)dt$$

onde f é uma função densidade de probabilidade de falha. A taxa instantânea de falha no momento imediatamente posterior a t_0 , dado que o sistema operava corretamente em t_0 é, então,

$$Z(t_0) = f(t_0) / R(t_0)$$

Finalmente, o tempo médio até a falha (MTTF, *Mean Time To Failure*) é

$$MTTF = \int_0^{\infty} tf(t)dt$$

Se a taxa de falhas λ_t for constante durante o intervalo de interesse, dispensa-se o subscrito e tem-se $Z(t) = \lambda$, $MTTF = 1/\lambda$ e $R(t) = e^{-\lambda t}$. Tal é a notação mais frequentemente utilizada neste trabalho.

3.8.1 Cálculo de confiabilidade e disponibilidade para redes sustentáveis

Até o trabalho de Amaral (2012), as abordagens para cálculo de confiabilidade e disponibilidade não consideravam a demora dos dispositivos em acordar (Lam e Li, 1986; Altıparmak, 2003; e He e Qi, 2008), considerando os equipamentos indisponíveis apenas em caso de falha. Consequentemente, a rede era tipicamente considerada estática, ou seja, o cálculo era efetuado sempre para a mesma topologia (Green, 2009; Yeh, 2010; e Lin, 2010). Em (Amaral, 2012), foi levado em conta o lapso temporal para acordar um dispositivo e a rede tornar-se estável; lapso o qual depende do sistema operacional, de alocações de memória, de outras atividades de comutação de estado energético, do protocolo de roteamento, etc. Amaral (2012) apresentou um método para cálculo de confiabilidade e disponibilidade em redes sustentáveis baseado em dois passos. Cada passo utiliza um algoritmo próprio. Os algoritmos são baseados no modelo de Márkov e no método dos conjuntos conectores ou dos conjuntos desconectores (Vejam-se as Seções 3.7 e 3.9). Entretanto, em (Amaral, 2012) não foi apresentado exatamente a que disposição de equipamentos corresponde o modelo lá discutido, nem foi apresentado como se obterem soluções analíticas para tal modelo. A presente dissertação realiza essas apresentações e mostra como se obter uma versão mais eficiente de um dos algoritmos, analisando sua complexidade. Os modelos de Márkov, e todo o método, também são estendidos.

3.9 Método dos conjuntos conectores ou dos conjuntos desconectores

A Seção 3.7 apontou que se utiliza o modelo de Márkov para, a partir de taxas de transição entre estados de um sistema, obter-se como calcular a confiabilidade ou disponibilidade desse sistema. O cálculo é preditivo e relativo a uma duração temporal, pois utiliza-se-o para prever quais as chances do elemento (nesta dissertação, enlace ou equipamento) estar funcionando daqui a alguns instantes (disponibilidade), ou quais as chances dele perdurar por certo tempo sem falhar (confiabilidade).

Ao contrário, o uso que será mostrado dos métodos do conjunto conector e do conjunto desconector tem natureza atemporal. Eles serão utilizados para, a partir da confiabilidade ou disponibilidade de componentes de um sistema, obter-se a confiabilidade desse sistema como um todo. Não raro, um componente pode ser ele mesmo um sistema. Essa situação é explorada pelo método discutido nesta seção. Ele faz uma síntese entre a natureza temporal do modelo de Márkov com a natureza atemporal do método do conjunto conector. Primeiramente, o modelo de Márkov será usado no cálculo da confiabilidade de um tipo pequeno de sistema, composto por conexões ou equipamentos redundantes e associados a taxas de falha. Em seguida, os valores da confiabilidade dos pequenos sistemas que constituem uma rede maior serão utilizados no método do conjunto conector para se obter a confiabilidade ou disponibilidade dessa rede maior.

Seja o sistema S (nesta dissertação, uma rede ou parte dela) composto pelos elementos $E_{1...e}$ (ou seja, há e elementos). Seja $R_{s,t}$ a confiabilidade de S relativa ao instante de tempo t , e $R_{1,t}..R_{e,t}$ a confiabilidade dos elementos componentes (nesta dissertação, enlaces ou equipamentos como roteadores ou comutadores) $E_{1...e}$, respectivamente, relativa ao mesmo instante de tempo t . Se a operação apropriada de S depende de todos os componentes estarem operando apropriadamente, então $R_{s,t} = \prod_{x=1}^{x=e} R_{x,t}$.

Agora suponha que para operar apropriadamente o sistema S não precise que todos os elementos componentes estejam operando apropriadamente. Nesse caso, o sistema precisa que apenas algum subconjunto C do total de elementos esteja funcionando. Ainda, pode haver outros tais subconjuntos. Cada um desses é um *conjunto conector*⁷. A união de dois conjuntos conectores é também um conjunto conector. Um elemento pode integrar mais de um conjunto conector. Nota-se que dois conjuntos conectores que tenham algum elemento em comum não são independentes quanto à probabilidade de estarem operando adequadamente ao mesmo tempo. Justamente essa falta de independência é explorada pelo método dos conjuntos conectores ou desconectores. Entretanto, importantes para o método são apenas os conjuntos conectores mínimos, ou seja, aqueles que não contêm como subconjunto algum outro conjunto conector.

De modo análogo, pode-se definir o conceito de conjunto desconector: um conjunto de elementos de um sistema é dito desconector quando o funcionamento inadequado de todos esses seus elementos implica o funcionamento inadequado do sistema. Um conjunto desconector é dito mínimo quando ele não contém nenhum subconjunto que seja um conjunto desconector.

Do exposto, vê-se que, no caso da operação apropriada do sistema S depender de todos os seus e componentes estarem operando apropriadamente, tem-se apenas um conjunto conector, $C_1 = \{E_1, E_2, \dots, E_e\}$, que também é mínimo. Ainda, há e conjuntos desconectores mínimos, cada qual correspondendo a cada um dos elementos, ou seja, $D_i = \{E_i\}$, com $1 \leq i \leq e$.

Definidos os conjuntos conectores mínimos ou desconectores mínimos pode-se proceder com o cálculo da confiabilidade do sistema, efetuado de modo combinatório (Shooman, 2001). A confiabilidade do sistema relativa ao instante t é probabilidade de pelo menos um conjunto conector mínimo operar

⁷ A terminologia fica mais clara se se pensar o sistema S como uma rede metropolitana, caso no qual os componentes $E_{1..e}$ são os roteadores (ou enlaces), que deve manter, digamos, a Cidade Universitária da Universidade de São Paulo conectada à Universidade Estadual de Campinas.

adequadamente até t , ou seja, é a união das probabilidades de operação adequada de cada um desses conjuntos.

Como os conjuntos envolvidos no cálculo podem não ser todos independentes, aplica-se o princípio da inclusão e exclusão, da teoria de conjuntos. Assim, seja $E_{x,t}$ o evento de x operar adequadamente até o instante t , e sejam $C_{1...n}$ os n conjuntos conectores do sistema S . A probabilidade $P(E_{S,t})$ de acontecer o evento $E_{S,t}$ é

$$P(E_{S,t}) = P(C_1 \text{ ou } C_2 \text{ ou } \dots \text{ ou } C_n) =$$

$$= \sum_1^n P(E_{C_i}) - \sum_{1 \leq i < j \leq n} P(E_{C_i} \cap E_{C_j}) + \sum_{1 \leq i < j < k \leq n} P(E_{C_i} \cap E_{C_j} \cap E_{C_k}) - \dots$$

(Equação 3.2)

Um exemplo do emprego da Equação 3.2 no contexto de redes é utilizá-la no cálculo da probabilidade de dois clientes de uma rede comunicarem-se. Nessa situação, cada conjunto conector corresponde a um caminho que permita aos dois clientes comunicarem-se, podendo haver dispositivos que pertençam a mais de um caminho.

O método dos conjuntos conectores é utilizado, nesta dissertação, na definição do método para avaliação da confiabilidade ou disponibilidade de redes sustentáveis (Capítulo 5). Alternativamente, seria possível a utilização do método dos conjuntos desconectores, análogo.

3.10 Considerações do capítulo

Este capítulo elencou iniciativas relacionadas à aplicação de sustentabilidade em redes de computadores e apresentou um embasamento teórico para o desenvolvimento de métodos de avaliação de sistemas de gerenciamento

energeticamente eficiente de redes. Foram apresentados também conceitos basais para a implementação de um sistema de gerenciamento orientado por política de eficiência energética e para a implementação de um ambiente de experimentação para tal sistema.

Há um esforço da IETF para determinar um conjunto de informações relacionadas a métricas de consumo de energia dos equipamentos de rede, os quais deverão fornecer informações sobre em que estado se encontram e permitir também que se determine em qual estado eles devem estar. Há também abordagens que consistem no desligamento de subsistemas de dispositivos. As técnicas apresentadas nas Seções 3.1 e 3.2 mostram um cenário em que há constante mudança de estado dos dispositivos. Em especial, a Seção 3.1 discute sistemas de gerenciamento que, através de seus algoritmos de roteamento, criam oportunidades para que outras técnicas sustentáveis atuem; por exemplo, o ElasticTree abre espaço para que roteadores ou comutadores entrem em modos ACPI de baixo consumo energético.

As Seções 3.3 – 3.5 dissertaram sobre técnicas que serão empregadas no sistema de gerenciamento apresentado e avaliado nos Capítulos 6 e 8. O trabalho de Chabarek e Barford (2011) foi especialmente útil para inspiração sobre novas extensões a ambientes de experimentação. Finalmente, as Seções 3.6 – 3.9 apresentaram o embasamento teórico dos Capítulos 4 e 5, onde se desenvolverão métodos para avaliação de soluções energeticamente eficientes em redes de computadores.

4 Método para avaliação do tempo de vida de equipamentos

Este capítulo discute como um sistema de gerenciamento de redes voltado a eficiência energética pode impactar os equipamentos previamente existentes na rede e como novos equipamentos sofrem impactos diferentes se suas funcionalidades verdes são utilizadas. Será apresentado como essas técnicas afetam o dispositivo internamente, modificando seu tempo médio até a falha (MTTF, *Mean Time To Failure*), ao passo que o Capítulo 5 mostra como utilizar o MTTF e outras grandezas no cálculo da confiabilidade de pequenos sistemas de uma rede (usualmente de dois a três dispositivos) e da rede toda. Após a apresentação dos principais mecanismos de falha de processadores e ventoinhas, será mostrado um método para avaliar se a solução verde adotada para uma rede é vantajosa economicamente.

4.1 Visão geral sobre a questão do tempo de vida

Há várias técnicas, muitas ainda em fase de desenvolvimento ou validação, que visam a melhorar a eficiência energética dos sistemas da área de TIC (Tecnologia da Informação e Comunicação). Os esforços mais proeminentes, de acordo com (Bolla, 2011b) e (Lee, 2011), são a) escalonamento de frequência de operação de equipamentos de acordo com a demanda, b) adaptação do modo energético de roteadores ou comutadores de acordo com a demanda de tráfego, e uma c) engenharia de tráfego mais eficiente, conforme já explicado no Capítulo 3. Uma variação de b) é colocar dispositivos para dormir de acordo com a demanda do tráfego, uma decisão que é melhor tomada após consulta aos modelos de energia do equipamento. Colocar um dispositivo em estado dormente afeta o estado da rede. Tal procedimento pode diminuir razoavelmente o consumo de energia elétrica da rede, mas pode implicar em comprometer seu desempenho e

confiabilidade, como mostrado em (Januário, 2012), (Gunaratne, 2008), e (Amaral, 2012). Ainda, ligar e desligar consuetudinariamente equipamentos afeta o tempo de vida esperado do *hardware*, como apresenta este capítulo. Assim, determinar o tempo de vida de equipamentos submetidos a mecanismos de eficiência energética passa a ser importante. Economicamente, esse tipo de informação não é útil apenas para quem usa os equipamentos, mas também para quem os vende. Com ela, fabricantes podem calcular os melhores prazos para as garantias fornecidas, de modo a maximizar o lucro.

Com o auxílio de abordagens probabilísticas, podem-se realizar estimativas do tempo de vida esperado desses equipamentos. O aparato probabilístico tipicamente se baseia em distribuições exponenciais, regidas por um parâmetro que indica uma taxa de falha (Neto e Cymbalista, 2006). Nesse caso, dada uma população de equipamentos semelhantes, o inverso do tempo de vida esperado dos equipamentos corresponde à taxa de falha observada na população.

Como exemplo de interesse para esse trabalho, pode-se citar o tempo de vida esperado das ventoinhas, comumente encontradas em sistemas de refrigeração de equipamentos, que é relacionado à temperatura (Jin, 2011), (Sidharth e Sundaram, 2004), (Tian, 2006) e a sua velocidade (Jin, 2011).

Apontou-se na Seção 2.4 que a ventoinha e a fonte de alimentação são os componentes que mais falham nos dispositivos de interesse. Entretanto, isso não significa que partes eletrônicas devam ser deixadas de lado, como observado em (Srinivasan, 2004). Os erros que podem ocorrer nas partes eletrônicas são relacionados ao desgaste do processador, causado após algum tempo devido à operação, mesmo que esta ocorra dentro das condições especificadas nos termos de garantia. Srinivasan (2004) expõe que, para lidar com confiabilidade e disponibilidade de processadores, é possível adaptar o funcionamento do processador como resposta a mudanças na sua carga de trabalho. Com isso, é possível assegurar-se um tempo de vida desejado. Srinivasan afirma que o processador sofre dois tipos de ciclos térmicos: a) ciclos térmicos grandes, que

ocorrem com baixa frequência, como quando se energizando ou desenargizando o processador para ligá-lo ou desligá-lo, ou como quando ele entra em modo dormiente ou desperta, e b) ciclos térmicos curtos, muito mais frequentes, devidos ao comportamento da carga do processador ou a mudança de contexto do processador.

Como consequência de estratégias voltadas a eficiência energética, como no caso da estratégia de beneficiar-se de diferentes estados energéticos, os ciclos térmicos grandes, antes infrequentes, passam a ser mais frequentes, tornando cogente uma análise dos possíveis impactos. Destarte, avaliar como a troca entre estados energéticos causa variação no MTTF do processador e do sistema de refrigeração associado é importante para a decisão sobre qual dispositivo pode ter seu estado energético alterado. O cálculo do MTTF é baseado nas condições de operação esperadas que o equipamento encontrará no seu campo de atuação. As condições são carga média, temperatura e voltagem, entre outras. Dentre as partes eletrônicas, o processador é aquela que sofre mais estresse. Dentre as partes mecânicas, a ventoinha é o dispositivo mais propenso a falhas. A fonte de alimentação está fora do escopo desta dissertação. Esta foca na avaliação do tempo de vida esperado de processadores e do sistema de refrigeração, pois tais partes podem ter o estado energético ou frequência de operação gerenciados diretamente, pertencendo, portanto, a um cenário mais rico.

Este capítulo se propõe a mostrar um método para avaliar a viabilidade de se colocarem equipamentos em modo dormiente ou acordá-los, com respeito ao tempo de vida esperado do próprio equipamento. Para tanto, as estratégias de eficiência energética que alteram o estado energético dos equipamentos de uma rede devem considerar o fato de poderem alterar o tempo de vida esperado desses equipamentos durante sua operação.

O restante deste capítulo está organizado assim: A Seção 4.2 disserta sobre o estado da arte relacionado a impactos no tempo de vida de processadores, e a

Seção 4.3, no tempo de vida das ventoinhas. A Seção 4.4 introduz um método de avaliação desses impactos e a Seção 4.5 conclui este capítulo.

4.2 Tempo de vida dos processadores

Os processadores compreendem as partes eletrônicas de interesse dos equipamentos discutidos, roteadores e comutadores. Foi apontado que o processador falha menos frequentemente que a ventoinha e a fonte de alimentação (Seção 2.3). Não obstante, não apenas os processadores mudam o ambiente físico (por exemplo, aquecendo-o), como também operá-los de maneira diferente da usual poderia afetá-los tão drasticamente que seu tempo de vida poderia tornar-se comparável ao dos demais componentes. Também poderia haver o caso de alguma condição de operação impelir o sistema operacional a subitamente desligar o dispositivo. Essa situação é obviamente indesejada.

As falhas dos processadores podem ser classificadas em duas categorias: erros leves e erros pesados⁸ (Rosing, 2007). Erros leves são falhas transientes ou distúrbios de um único evento. Eles ocorrem durante a execução do processador e são devidos a radiação externa ou a ruído elétrico; não são consequência de defeitos de projeto ou manufatura. Embora erros leves possam causar erro de computação ou corrupção de dado, eles não danificam o processador e não são considerados no cômputo da confiabilidade do equipamento (Srinivasan, 2004). Erros pesados advêm de “defeitos na parte de silício ou de metalização da cápsula do processador, sendo usualmente permanentes uma vez manifestos” (Srinivasan, 2004). Esse tipo de erro ocasiona falha permanente no processador (Rosing, 2007). Porque erros pesados resultam em falha permanente, o tempo de vida esperado nesse caso é inversamente proporcional à taxa de ocorrência desse

⁸ Neste trabalho, “erro leve” e “erro pesado” traduzem respectivamente o que na literatura de língua inglesa é “*soft error*” e “*hard error*”.

tipo de erro. Portanto, esta taxa diretamente determina o tempo de vida dos equipamentos em pauta.

Erros pesados podem ser subdivididos em falhas extrínsecas e falhas intrínsecas (Pech, 1998). Falhas extrínsecas ocorrem numa taxa que decresce com o tempo, sendo ocasionadas por defeitos de manufatura. Falhas intrínsecas estão relacionadas à degenerescência do processador, sendo ocasionadas durante a operação, mesmo que a operação se dê de acordo com as condições especificadas pelo fabricante. Falhas intrínsecas dependem do tipo de material usado na fabricação do processador, estando relacionadas ao projeto deste e ao modo como este é encapsulado na bolacha de sílica. Vê-se que, não havendo nem erros relacionados ao processo de fabricação nem ao projeto e uso, então todas as falhas restantes são ocasionadas por falhas intrínsecas (Klutke, 2003).

O tempo de vida do processador está diretamente relacionado à temperatura de operação, donde se esperam problemas caso a temperatura do processador seja drasticamente alterada pela nova operação sustentável. Usualmente, a relação entre as taxas de falhas de um processador em diferentes temperaturas é expressa numa forma semelhante à equação de Arrhenius. Esse tipo de equação modela a dependência observada que a taxa de uma reação química tem em relação à temperatura ou a mudanças de temperatura (JEDEC, 2006). Srinivasan (2004) asseverou que os mecanismos críticos de falhas intrínsecas são a eletromigração, a migração de estresse, a ruptura da porta de óxido ou TDDB (*Time Dependent Dielectric Breakdown*, ruptura dielétrica dependente do tempo), e os ciclos térmicos. Rosing (2007) foca na eletromigração, no TDDB e nos ciclos térmicos, sem nem citar a migração por estresse. Descrevem-se a seguir os quatro mecanismos.

4.2.1 Eletromigração

A eletromigração é consequência da depleção de metal de alguma região do condutor e seu acúmulo em outras regiões (Dasgupta e Karri, 1996). Esse fenômeno ocorre devido a transferência de energia, causada pela corrente de elétrons, para as junções de metal. Se a corrente for suficientemente alta, o mecanismo é ativado. O resultado é uma aparente abertura de circuito na região donde o metal sai (Black, 1969). O modelo de Black para o MTTF devido à eletromigração é dado pela Equação 4.1.

$$MTTF_{EM} = A_0(J - J_{crit})^{-n} e^{\frac{E_a}{kT}} \quad (\text{Equação 4.1})$$

Na Equação 4.1, tem-se:

- A_0 é uma constante determinada empiricamente;
- J é a densidade de corrente que atravessa as junções metálicas, devendo ser maior que J_{crit} para produzir falha;
- J_{crit} é a densidade de corrente crítica, acima da qual o mecanismo de eletromigração é ativado;
- E_a é a energia de ativação da reação, cujo valor típico varia de 0,5 a 0,9 (JEDEC, 2006);
- k é a constante de Boltzmann;
- T é a temperatura absoluta; e
- n é outra constante empírica, cujo valor depende do tipo das junções metálicas. Valores típicos de n vão de 1 a 2 (Srinivasan, 2003). Para ligas de alumínio, E_a e n valem respectivamente 0,7 e 2. (Rosing, 2007).

Em uma operação sustentável em que os equipamentos passem mais tempo em modos de baixo consumo, a temperatura média no processador é menor do que seria se ele estivesse sempre ligado. Com uma temperatura média menor, o expoente do número de Euler aumenta, bem como o $MTTF_{EM}$.

4.2.2 Migração por estresse

A migração por estresse é ocasionada pelo movimento de átomos de metal, que por sua vez é ocasionado por estresse mecânico (JEDEC, 2006). Uma das modalidades desse mecanismo é termomecânica. Nesse caso, taxas de expansão térmica diferentes em diferentes regiões da bolacha de sílica levam a estresse. A Equação 4.2 relaciona a migração por estresse à temperatura.

$$MTTF_{ME} = A_0 |T - T_0|^{-n} e^{\frac{E_a}{kT}} \quad (\text{Equação 4.2})$$

Na Equação 4.2, tem-se:

- T é a temperatura absoluta;
- T_0 é a temperatura sob a qual o encapsulamento fora montado, portanto sob a qual não há estresse. T_0 é usualmente muito maior que T, podendo chegar, segundo (Srinivasan, 2004), a 500K para algumas monstagens;
- A_0 é uma constante determinada empiricamente, dependente do material;
- n é uma constante empírica, também dependente do material; e
- E_a é a energia de ativação da reação, também determinada empiricamente.

De modo análogo ao que ocorre para a Equação 4.1, uma temperatura média menor, esperada na operação sustentável, faz o expoente do número de Euler aumentar, tendendo a acrescer o $MTTF_{ME}$. De modo contrário, a componente $|T - T_0|^{-n}$ tende a fazer o $MTTF_{ME}$ diminuir, pois a temperatura de montagem T_0 é tipicamente muito maior que a temperatura de operação. Rosing (2007), entretanto, ensina que o efeito geral de uma menor temperatura de operação ainda é o aumento do $MTTF_{ME}$.

4.2.3 Ruptura dielétrica dependente do tempo

Dielétricos são usados em microeletrônica devido a sua propriedade isoladora. A ruptura do dielétrico é a formação de um caminho de baixa resistência condutiva através do dielétrico, formação a qual compromete as propriedades isoladoras, ocasionando falhas irreversíveis no equipamento. A confiabilidade da porta de óxido depende da temperatura, da tensão aplicada à porta, e do campo elétrico nesta. Nesse mecanismo de falha, a temperatura desempenha uma influência mais que exponencial (Srinivasan, 2004). Wu (2002) modela o MTTF devido à ruptura dielétrica dependente do tempo através da Equação 4.3.

$$MTTF_{TDD} \propto \left(\frac{1}{V}\right)^{a-bT} e^{\frac{(X+\frac{Y}{T}+Z.T)}{kT}} \quad (\text{Equação 4.3})$$

Na Equação 4.3, tem-se:

- a,b, X, Y e Z são determinadas empiricamente. Para portas de óxido ultrafinas, Wu (2002) e Srinivasan (2004) utilizam esses valores como a = 78, b = -0,081, X = 0,759ev, Y = -66,8evK, e z = -8,37e - 4ev/K;
- T é a temperatura absoluta; e
- V é a voltagem na porta de óxido.

4.2.4 Ciclos térmicos

A variação de temperatura também causa fadiga. O acúmulo dessa fadiga toda vez que há um ciclo térmico eventualmente pode ocasionar falha. Podem-se dividir os ciclos térmicos em duas categorias: longos e curtos. Ciclos da primeira categoria ocorrem menos frequentemente; algumas vezes ao dia, quando muito. Estes são relacionados a energizar ou desenergizar o equipamento para ligá-lo ou

desligá-lo, ou entrar em modo dormente ou despertar. Ciclos térmicos curtos ocorrem muito mais frequentemente, algumas vezes por segundo, sendo devidos a mudanças na intensidade da carga ou, quando o processador permite, a troca de contexto⁹. Ciclos térmicos curtos não são ainda muito bem estudados pela comunidade, de sorte que ainda não há para eles modelos validados (Srinivasan, 2004). Ciclos térmicos longos são regidos pela equação de Coffin-Manson (JEDEC, 2006), mostrada aqui na Equação 4.4:

$$N = C_0(\Delta T)^{-q} \quad (\text{Equação 4.4})$$

Na Equação 4.4, tem-se:

- N é o número (médio) de ciclos até a falha (MCTF);
- C_0 é uma constante dependente do material;
- ΔT é a variação de temperatura; e
- q é uma constante. Srinivasan (2004) utiliza o valor 2,35 para a constante q de processadores.

Dado um tipo de ciclo térmico caracterizado pela diferença de temperatura ΔT_1 e pela frequência esperada de ocorrência f_1 , vê-se que o MTTF implicado pelo MCTF (*Mean Cycles To Failure*) é a divisão desta quantia pela frequência f_1 (Equação 4.5).

$$MTTF_{TC} = \frac{MCTF}{f} = \frac{N}{f} \quad (\text{Equação 4.5})$$

⁹ A troca de contexto, em um processador, dá-se quando este para de executar uma tarefa, à qual ele possivelmente retornará, e passa a executar outra. Nesse tipo de operação, o processador armazena em memória o estado da tarefa interrompida e carrega o estado da tarefa (re)iniciada.

Dado o MCTF, relativo a ciclos de certa variação de temperatura, de um processador, a confiabilidade R_c deste sobreviver após x ciclos é dada pela Equação 4.6:

$$R_c(x) = P(\text{ciclos} > x) = e^{-\frac{x}{MCTF}} \quad (\text{Equação 4.6})$$

4.3 Tempo de vida das ventoinhas

As ventoinhas compreendem a parte mecânica de interesse dos equipamentos discutidos nesta dissertação. Elas são compostas por invólucro, rolamento, lâminas, eixo, motor, partes eletrônicas, etc. A confiabilidade da ventoinha é um parâmetro crítico no sistema de refrigeração de produtos eletrônicos (Jin, 2011). As falhas da ventoinha podem ser classificadas em mecânicas ou elétricas. Segundo Wilson e Wilkinson (1975), as falhas mecânicas no rolamento correspondem a 80,85% das falhas totais, ao passo que falhas eletrônicas correspondem a apenas 16,55%. Os outros tipos de falha mecânica correspondem a apenas 2,60%.

Como no caso dos processadores, o tempo de vida de ventoinhas também é intimamente relacionado à temperatura, como apontam Jin (2011), Sidharth e Sundaram (2004) e Tian (2006). Também os ciclos térmicos influenciam significativamente o tempo de vida, junto com o tipo de rolamento, a lubrificação e a velocidade de operação. Entretanto, pequenas durações de alta carga não reduzem significativamente o tempo de vida da estrutura de rolamento (Wilson e Wilkinson, 1975). Quase todas as falhas na estrutura de rolamento podem ser rastreadas até uma correspondente degradação da lubrificação (Jin, 2011). Em seu trabalho, Wilson e Wilkinson mostram que a vida do lubrificante decresce pela metade a cada aumento de 10°C a 15°C na temperatura ambiente.

A temperatura afeta a ventoinha principalmente por catalisar a degenerescência de sua estrutura, um efeito que é devido à constituição ou posicionamento, a danos na lubrificação, e a outros fatores. Modelos baseados na equação de Arrhenius são tipicamente usados na análise desse efeito. Wilson e Wilkinson (1975) analisaram um banco de dados sobre falhas com mais de 2000 ventoinhas eletrônicas e determinaram equações para os tipos de falha mais representativos: no rolamento e na bobina. Os mesmos autores também atentaram aos ciclos térmicos. Descrevem-se a seguir os principais mecanismos de falha das ventoinhas.

4.3.1 Tempo de vida do rolamento

O tempo de vida do rolamento da ventoinha é função do tempo de vida da graxa, para temperaturas maiores que 77°C. Pode-se-o descrever empregando-se a distribuição de Weibull com parâmetro de forma $\beta = 3$ e parâmetro de escala α , também chamado de vida característica, dado pela Equação 4.7.

$$\log \alpha = \frac{2450}{T} - K_g - q \frac{DN}{DN_L} \quad (\text{Equação 4.7})$$

Na Equação 4.7, tem-se:

- T é a temperatura do rolamento (usa-se um valor de 40°C mais elevado que a temperatura ambiente quando o aumento de temperatura ocasionado pela bobina é desconhecido);
- K_g é uma constante relativa à graxa (usa-se o valor de 2 quando o diâmetro da estrutura é desconhecido);
- q mede o grau de qualidade da ventoinha (usa-se $q = 0,26$ para dispositivos militares e $q = 0,95$ para dispositivos comerciais);

- DN é o produto entre o diâmetro interno da estrutura (medido em milímetros) e a velocidade (medida em rotações por minuto);
- e DN_L é o valor limite para DN. Quando a temperatura do suporte é inferior a 77°C e maior que 37°C, o tempo de vida deve ser predito pela Equação 4.8:

$$\log \alpha = -\frac{2450}{T} + B \quad (\text{Equação 4.8})$$

Na Equação 4.8, $B = 12$ para ventoinhas de qualidade comercial e $B = 12,3$ se a qualidade é militar. Um segundo cuidado deve ser tomado. Após aplicada a Equação 4.8, se o tempo de vida previsto exceder o valor limite de 80000 horas para qualidade comercial, ou 120000 para militar, o correspondente valor limite deve ser adotado, não o valor resultante da equação. Wilson e Wilkinson (1975) não consideram efeitos que seriam ocasionados pela velocidade, alegando que, a temperaturas baixas, os efeitos haveriam de se negar mutuamente.

As Equações 4.7 e 4.8 predizem o tempo de vida do suporte com base em apenas uma temperatura. Sem embargo, podem-se postular outros níveis de velocidade e calcular tempos de vida distintos, cada qual correspondendo a cada nível. A análise obteria para o suporte um MTTF relacionado a diferentes velocidades de operação. Uma análise semelhante é apresentada na Seção 4.3.3. Incluir no cálculo diferentes velocidades é mister quando se analisando futuros sistemas de gerenciamento energeticamente eficiente que precisem gerenciar o sistema de refrigeração dos equipamentos.

4.3.2 Tempo de vida da bobina

O tempo de vida da bobina é função da degradação causada na isolação elétrica. Novamente, uma distribuição de Weibull é apropriada. A vida característica da bobina pode ser predita pela Equação 4.9. Tal valor deve ser usado como

parâmetro de escala de uma distribuição de Weibull de parâmetro de forma $\beta = 1,76$.

$$\log \alpha = \frac{4170}{T} - A \quad (\text{Equação 4.9})$$

Na Equação 4.9, tem-se:

- T é a temperatura absoluta do ponto mais quente da bobina. A temperatura T é a soma da temperatura ambiente com o aumento devido à bobina, com mais 10K. Caso não seja conhecido o aumento devido à bobina, Wilson e Wilkinson sugerem assumi-lo correspondente a 50°C;
- e A é a constante de isolamento, adimensional.

4.3.3 Ciclos térmicos

Quando a operação é conduzida ciclicamente passando por diferentes temperaturas, o tempo de vida pode ser predito pelas Equações 4.10 e 4.11.

$$n \left(\frac{h_1}{L_1} + \frac{h_2}{L_2} + \dots + \frac{h_m}{L_m} \right) = 0,5 \quad (\text{Equação 4.10})$$

$$MTTF = n(h_1 + h_2 + h_m) \quad (\text{Equação 4.11})$$

Nas Equações 4.10 e 4.11, tem-se:

- h_m é o tempo gasto por ciclo na m-ésima temperatura do ciclo;
- L_m é o tempo de vida esperado caso a operação se desse apenas sob a m-ésima temperatura; e

- n é o número médio de ciclos até a falha.

Na Equação 4.10 recomenda-se a utilização do valor 0,5, ao invés do esperado valor unitário, porque “qualquer extrapolação dos dados existentes [na pesquisa de campo deles] é pura especulação” (Wilson e Wilkinson, 1975), sendo sugerida a cautela de se diminuir à metade o número esperado de ciclos.

4.4 Método de análise

O método de análise parte do modelo de soma de taxa de falhas (SOFR, *Sum-Of-Failure Rates*). Esse modelo é útil para determinar o MTTF de um sistema, como apontado em (Rosing, 2007). O modelo SOFR é aceito pela indústria de semicondutores no cálculo de confiabilidade. Ele trata a falha de cada estrutura do sistema como independente das falhas das demais estruturas. Wilson e Wilkinson (1975) empregam o mesmo modelo para prever o tempo de vida da ventoinha, sugerindo que se determine o MTTF de todas as partes constituintes do equipamento. A confiabilidade do equipamento é o produto da confiabilidade dos mecanismos. Cada mecanismo age independentemente dos demais, e o primeiro a falhar faz o equipamento falhar.

Seja o processador P . Pelo modelo SOFR, sua taxa de falha é $\lambda_P = \lambda_{EM} + \lambda_{ME} + \lambda_{TDDDB} + \lambda_{TC}$. A confiabilidade fica então $R_P(t) = e^{-(\lambda_{EM} + \lambda_{SM} + \lambda_{TDDDB} + \lambda_{TC})t} = R_{EM}(t) \times R_{ME}(t) \times R_{TDDDB}(t) \times R_{TC}(t)$. Assim, se a operação sustentável acarreta uma diminuição de $x\%$ na confiabilidade de um mecanismo, sem alterar a dos demais, então a confiabilidade do processador diminui também em $x\%$.

Rosing (2007) mostra que o MTTF do mecanismo de falhas ocasionado por ciclos térmicos é o que mais decresce em função de uma operação sustentável. Não só é o que mais decresce, mas é o que mais rapidamente é alterado, como mostra a

Figura 2. Essa figura mostra o compromisso entre diferentes mecanismos de falha e ganhos de energia, para a tecnologia 95-nm de um fabricante parceiro de Rosing (2007).

Figura 2 - MTTF relativo a um núcleo de processador de tecnologia 95-nm, para diferentes mecanismos de falha, quando diferentes ganhos de energia são obtidos (Rosing, 2007)

Com as operações energeticamente eficientes de redes, espera-se que os equipamentos transitem entre diferentes estados energéticos. Assim, suponha-se que haja os estados energéticos E_1 , E_2 e E_3 e que os ciclos térmicos possam ser do tipo 1, com amplitude ΔT_1 , ocorrendo entre E_1 e E_2 , ou do tipo 2, com amplitude ΔT_2 , ocorrendo entre E_2 e E_3 . Nessas condições, dado que o processador estava funcionando antes do primeiro ciclo, a probabilidade dele estar funcionando após x_1 ciclos do tipo 1 é $P_1 = e^{-\frac{x_1}{MCTF_1}}$. Analogamente, se o processador está funcionando antes do primeiro ciclo do tipo 2, a probabilidade dele sobreviver a x_2 ciclos desse tipo é $P_2 = e^{-\frac{x_2}{MCTF_2}}$. Do mesmo modo, prova-se que a confiabilidade

de um período durante o qual o processador passa por x_1 ciclos do tipo 1 e x_2 ciclos do tipo 2 efetuados em qualquer ordem é dada pela Equação 4.12

$$R = e^{-\left(\frac{x_1}{MCTF_1} + \frac{x_2}{MCTF_2}\right)} \quad (\text{Equação 4.12})$$

É fácil ver que, se os ciclos são do mesmo tipo, então $\Delta T_1 = \Delta T_2 \leftrightarrow N_1 = N_2$ e vale a Equação 4.13 (com $MCTF = N_1 = N_2$), onde se tem a probabilidade do processador sobreviver após x_1+x_2 ciclos do mesmo tipo.

$$R = e^{-\left(\frac{x_1+x_2}{MCTF}\right)} \quad (\text{Equação 4.13})$$

A Equação 4.12 pode ser estendida para abranger vários tipos (várias amplitudes) de ciclos. Seu valor preditivo é de especial utilidade quando se dispõe de uma estimativa sobre a frequência de ocorrência dos ciclos, a qual pode ser obtida a partir da estimativa de ocorrência das transições entre estados energéticos. Por sua vez, esta ocorrência é função (i) das funcionalidades de eficiência energética aplicadas à rede (SSC ou ALR, por exemplo) ou de como estas são gerenciadas (como o sistema operacional do equipamento ou o sistema de gerenciamento da rede decide usar os modos ACPI, por exemplo) e (ii) do perfil de tráfego esperado nos equipamentos.

4.5 Considerações do capítulo

A temperatura de operação é o principal fator que altera o tempo médio até a falha dos equipamentos de rede, estando relacionada a mecanismos de falha nos processadores e nas ventoinhas. Nos processadores, os ciclos térmicos estão também relacionados a mecanismos de falha, de forma que, mesmo que a temperatura de operação não mude, desenergizar um equipamento e depois reenergizá-lo causa desgaste. Para as ventoinhas, quando a operação se dá ciclicamente a diferentes temperaturas, recomenda-se reduzir à metade o tempo

de vida calculado (Equação 4.10). Dos mecanismos de falha em roteadores, o causado por ciclos térmicos é o que mais se modifica durante uma operação sustentável que troque o nível de energização dos equipamentos. A partir de uma estimativa sobre quantas vezes o processador passará por cada tipo de ciclo térmico, pode-se estimar como seu tempo de vida foi afetado em relação a uma operação usual, na qual ele está sempre ligado. Essa informação pode auxiliar na decisão da viabilidade econômica ao se implantar uma solução verde.

5 Método para avaliação da confiabilidade e disponibilidade de redes sustentáveis

Para o cálculo da confiabilidade e disponibilidade de redes que utilizam técnicas de eficiência energética como a aplicação de modo dormente, foi proposto o método REASoN (Amaral, 2012). O autor propõe a utilização de modelos de Márkov e os métodos dos conjuntos conectores ou desconectores para calcular, respectivamente, a confiabilidade ou disponibilidade primeiramente de um elemento de rede, depois da rede toda. Este capítulo apresenta a que disposição de elementos de rede (Figura 4) correspondem alguns modelos de Márkov (Figura 5) e descreve extensões ao método REASoN.

5.1 Contexto do método proposto

Como discutido no Capítulo 3, algumas soluções para eficiência energética para redes de computadores colocam equipamentos inteiros ou apenas algumas de suas interfaces em modos de consumo de energia elétrica reduzido. Nesses modos, os equipamentos não se encontram completamente operacionais. Nesse tipo de solução verde, o ápice de redução de energia ocorre quando o equipamento todo é posto em modo dormente. Consoante observado alhures (Amaral, 2012), do ponto de vista do desempenho da rede, o modo dormente é similar a deixar o dispositivo em reserva¹⁰, ou seja, desligado.

A contraparte do método de reserva é o de reserva quente (*hot stand by*). Neste, o dispositivo permanece completamente operacional e útil, pronto para encaminhar pacotes ou desempenhar outras funções de roteamento tão logo seja necessário.

¹⁰ Optou-se por traduzir a expressão inglesa “(*cold*) stand by” como “em reserva”, seguindo (Neto e Cymbalista, 2006).

Ao contrário, conforme observado na Seção 3.2, um dispositivo em modo dormiente demoraria um tempo não negligenciável para tornar-se útil. Assim, do ponto de vista de desempenho, é preferível a uma operadora de rede utilizar sistemas com dispositivos apenas em reserva quente, ao invés de aplicar uma reserva normal.

Ora, acordar sobre o desempenho do serviço provido é uma constante nos acordos de nível de serviço (SLA, *Service Level Agreement*) que são travados entre uma operadora e seus clientes. Para uma operadora que deseje ser energeticamente eficiente, é então crucial precisar como o uso de sistemas em reserva faria a disponibilidade e a confiabilidade de sua rede variar em relação ao uso ordinário. Somente com esse tipo de informação a operadora pode tomar decisões bem embasadas sobre os efeitos que o tempo gasto no acordar de um equipamento, ou seja, na mudança entre modos energéticos, acarreta na qualidade do serviço prestado. Assim, a operadora pode avaliar quando e quais nós podem entrar em modo dormiente concomitantemente a uma perda aceitável de desempenho.

Justamente para avaliar essas questões foi desenvolvido o REASoN. Este é o método tomado por base por esta dissertação para cálculo de confiabilidade e disponibilidade em redes sustentáveis. Ele leva em conta o dinamismo previsto para essas redes e a demora em se acordarem os dispositivos. O REASoN combina duas técnicas bem conhecidas, nomeadamente a modelagem por Márkov e a teoria dos conjuntos conectores ou desconectores. Após a apresentação do REASoN, descreve-se sua nova geração, o Reasoning.

5.2 O método REASoN

O REASoN (*REliability and/or Availability evaluation for Sustainable Networking*, Avaliação de Confiabilidade ou Disponibilidade para Redes Sustentáveis) é um

método que intenciona medir os efeitos colaterais que o uso de mudanças de estado energético de equipamentos, mais especificadamente, o uso do estado dormente, causa na rede (Amaral, 2012). O método permite a predição da nova confiabilidade e disponibilidade da rede caso algum equipamento tenha seu estado energético modificado para o estado dormente. Um sistema de gerenciamento de redes pode utilizar a predição para decidir se uma adaptação dinâmica almejada em um dispositivo é aceitável ou não na perspectiva da qualidade de serviço.

Há dois passos no REASoN. O primeiro passo do REASoN “usa um modelo estendido de Márkov para calcular, para cada comutador ou roteador, a confiabilidade ou disponibilidade” (Amaral, 2012), considerando a demora para despertar e as conexões redundantes. O segundo passo é utilizar o método dos conjuntos conectores ou desconectores para, a partir dos resultados individualizados do primeiro passo, calcular a confiabilidade e a disponibilidade da rede como um todo. Este passo “considera como indisponível qualquer caminho composto [...] por algum roteador em reserva (ou inativo)” (Amaral, 2012).

A Figura 3 mostra o cerne do método REASoN. Ela mostra um modelo de Márkov que representa a probabilidade de um equipamento de rede estar “encaminhando pacotes de dados” (Amaral, 2012), ou seja, estar funcionando adequadamente. O cerne do método é modelar os efeitos do *cold standby* através da adição¹¹ de um estado de penalidade através do qual a demora em despertar do modo dormente é representada. Nos modelos da Figura 3, λ é a taxa de falha do dispositivo, α é o inverso do tempo que demora até que o dispositivo acorde e a rede estabilize, μ_1 é a taxa de reparo preventivo e μ_2 é a taxa de reparo após o sistema ter falhado. O estado 1S representa um roteador ativo e outro em reserva; *Delay* representa que o roteador antes ativo falhou e que o roteador antes em reserva está despertando; 01 representa o roteador falho e o outro já acordado, com a rede estabilizada; e 00

¹¹ Adição em relação a um modelo de Márkov, usual, que não modele a demora em despertar.

representa que os dois roteadores estão falhos. Como $\alpha \gg \lambda$, a taxa de falhas pode ser desconsiderada quando no estado *Delay*.

Figura 3 - Modelos de Márkov para confiabilidade e disponibilidade usados pelo REASoN, com taxa de cobertura de 100%. Título original: *Extended Markov model of a router with two connections in cold standby that considers the time to activate the standby* (Amaral, 2012)

O estado *Delay* representa uma penalidade, cuja consequência é que, nas palavras de Amaral (2012), “a probabilidade de estar nesse estado interfere na probabilidade da rede estar operando”. Ainda sobre o estado de penalidade, “a probabilidade de um dispositivo estar nesse estado”, prossegue Amaral, “é computada na probabilidade total do sistema estar nos estados não operacionais”.

Tomando por base a Figura 3, vêm as Equações 5.1 e 5.2:

$$P_{1S}(t + \Delta t) = P_{1S}(t) - \lambda \cdot \Delta t \cdot P_{1S}(t)$$

$$P_{Delay}(t + \Delta t) = P_{Delay}(t) - \alpha \cdot \Delta t \cdot P_{Delay}(t) + \lambda \cdot \Delta t \cdot P_{1S}(t)$$

$$P_{0S}(t + \Delta t) = P_{0S}(t) - \lambda \cdot \Delta t \cdot P_{0S}(t) + \alpha \cdot \Delta t \cdot P_{Delay}(t)$$

$$P_{00}(t + \Delta t) = P_{00}(t) + \lambda \cdot \Delta t \cdot P_{0S}(t)$$

$$R(t + \Delta t) = P_{1S}(t + \Delta t) + P_{0S}(t + \Delta t) = 1 - (P_{00}(t + \Delta t) + P_{Delay}(t + \Delta t))$$

(Equação 5.1)

$$P_{1S}(t + \Delta t) = P_{1S}(t)(1 - \lambda \cdot \Delta t) + \mu_1 \Delta t \cdot P_{0S}(t) + \mu_2 \cdot \Delta t \cdot P_{00}(t)$$

$$P_{Delay}(t + \Delta t) = P_{Delay}(t)(1 - \alpha \cdot \Delta t) + \lambda \cdot \Delta t \cdot P_{1S}(t)$$

$$P_{0S}(t + \Delta t) = P_{0S}(t)(1 - \lambda \cdot \Delta t + \mu_2 \cdot \Delta t) + \alpha \cdot \Delta t \cdot P_{Delay}(t)$$

$$P_{00}(t + \Delta t) = P_{00}(t)(1 - \mu_1 \Delta t) + \lambda \cdot \Delta t \cdot P_{0S}(t)$$

$$A(t + \Delta t) = P_{1S}(t + \Delta t) + P_{0S}(t + \Delta t) = 1 - (P_{00}(t + \Delta t) + P_{Delay}(t + \Delta t))$$

(Equação 5.2)

Dos modelos de Márkov obtêm-se as expressões para a probabilidade do sistema estar em cada estado no momento t (Equações 5.1 e 5.2). Essas equações são utilizadas no primeiro passo do REASoN. Elas correspondem a equações diferenciais, como aponta o termo Δt . Podem-se-as resolver analiticamente, mas essa pode ser uma tarefa muito complexa (Shooman, 2001), de forma que a solução analítica para a modelagem proposta pelo REASoN ainda não foi apresentada na literatura. Uma solução alternativa é através de iterações conforme o Algoritmo 5.1, extraído de (Amaral, 2012). O procedimento é semelhante a uma simulação, pois, dado o passo de tempo e a configuração inicial do sistema, simula-se qual a probabilidade do sistema estar em cada estado, passo após passo, até o instante desejado. As condições de contorno usuais, e também apropriadas ao problema em mãos, são de que no instante $t = 0$ o sistema se encontra com todos os elementos em perfeitas condições, ou seja, $P(0)_{1S} = 1$, $P(0)_{Delay} = P(0)_{0S} = P(0)_{00} = 0$. O resultado obtido pelo método iterativo depende do passo da iteração, pois quanto menor o passo, melhor o resultado. Daí o valor de alguma técnica que diminua a complexidade temporal das iterações, como a que será mostrada na Seção 5.3.2. A (in)equação abaixo demonstra como passos de diferentes magnitudes levam a resultados diferentes.

Na (in)Equação 5.3, $P_{e,\Delta T}(t)$ é a probabilidade do sistema estar no estado e , no tempo t , calculada com passo ΔT após sucessivas aplicações da Equação 5.1.

$$P_{1S,2}(2) = P_{1S,2}(0) \cdot (1 - \lambda \cdot 2) = 1 - 2\lambda \neq (1 - \lambda)^2 = P_{1S,1}(1) \cdot (1 - \lambda) = P_{1S,1}(2) \quad (\text{Equação 5.3})$$

De fato, se $x \neq y$, $P_{1S,x}(t) = P_{1S,y}(t)$ apenas para $\lambda = 0$. Assim, um estudo sobre qual valor de ΔT levaria a uma solução aceitável faz-se necessário, remanesecendo como trabalho futuro.

Algoritmo 5.1: Processo iterativo do REASoN para cálculo de $R(t)$ de um dispositivo

Laço de $t = 0$ a $t = t_{\text{fim}} - \Delta t$

Se $t = 0$ então

$$P_0(0) = 1$$

$$P_1(0) = 0$$

$$P_2(0) = 0$$

$$P_3(0) = 0$$

$$R(t) = 0$$

Senão

$$P_0(t + \Delta t) = P_0(t) - \lambda \cdot \Delta t \cdot P_0(t)$$

$$P_1(t + \Delta t) = P_1(t) - \alpha \cdot \Delta t \cdot P_1(t) + \lambda \cdot \Delta t \cdot P_0(t)$$

$$P_2(t + \Delta t) = P_2(t) - \lambda \cdot \Delta t \cdot P_2(t) + \alpha \cdot \Delta t \cdot P_1(t)$$

$$P_3(t + \Delta t) = P_3(t) + \lambda \cdot \Delta t \cdot P_2(t)$$

$$R(t + \Delta t) = P_0(t + \Delta t) + P_2(t + \Delta t)$$

Fim Se

$$t = t + \Delta t$$

Fim Laço

Retorna $R(t)$

O segundo passo do REASoN corresponde ao Algoritmo 5.2. Ele pode ser realizado pela lógica tanto dos conjuntos conectores, quanto dos conjuntos desconectores. Este passo corresponde a fazer uma síntese da confiabilidade ou disponibilidade individual dos equipamentos relativa ao instante t , obtendo a confiabilidade ou disponibilidade da rede como um todo, relativa ao mesmo instante.

Algoritmo 5.2: Processo iterativo do REASoN para cálculo de $R(t)$ da rede, utilizando o método dos conjuntos conectores

Dados

C conjunto dos conjuntos conectores

c cada elemento de C

$R_{x,t}$ confiabilidade de x (rede ou dispositivo) relativa ao instante t

t instante relativo ao qual se deseja a confiabilidade

Para cada c em C

Para cada dispositivo r em c

$R_{c,r} = \text{Algoritmo 5.1}(r, t)$

Fim para cada

Fim para cada

$R_{\text{rede},t} = \text{Equação 3.2}(C)$

Retorna $R_{\text{rede},t}$

Nota-se uma ineficiência no Algoritmo 5.2: embora o resultado seja o mesmo, a instrução do laço interno (ou seja, todo o Algoritmo 5.1) é executada mais de uma vez para o mesmo equipamento, caso este integre mais de um conjunto conector. Uma alternativa seria executar-se o Algoritmo 5.1 para todos os equipamentos da rede, a despeito de se saber a qual conjunto conector cada um pertence, e posteriormente efetuar o cálculo da Equação 3.2. Entretanto, esta alternativa poderia efetuar o desnecessário cálculo do Algoritmo 5.1 para um dispositivo que não integrasse nenhum conjunto conector.

5.3 Método proposto – do REASoN ao Reasoning

Esta seção apresenta o método Reasoning (*REASoN's New Generation*, Nova Geração do REASoN). Reasoning é um método para avaliação da disponibilidade e confiabilidade de redes sustentáveis derivado do REASoN. As semelhanças e diferenças encontram-se resumidas no Quadro 1 e são abordadas no restante deste capítulo.

Como herança do método REASoN, o Reasoning avalia a confiabilidade e disponibilidade em dois passos. Entretanto, o primeiro passo do REASoN “usa um modelo estendido de Márkov para calcular, para **cada comutador ou roteador**, a confiabilidade ou disponibilidade” (Amaral, 2012; o grifo é desta dissertação), considerando a demora para despertar e as conexões redundantes. Esse procedimento deve ser aprimorado, pois o efeito do estado dormiente é avaliado pela interação do elemento dormiente com os adjacentes, como ficará patente ao se discutirem as Figura 4 e Figura 5, de modo que o efeito da dormência não possa ser isolado para apenas um elemento. O primeiro passo do Reasoning é, assim, não usar um modelo de Márkov para calcular a confiabilidade ou disponibilidade de cada roteador (elemento) da rede (sistema), mas sim de cada subsistema, ou seja, de cada conjunto conveniente de elementos da rede onde seja possível isolar o efeito da dormência de um elemento.

Método / Característica	REASoN	Reasoning
Permite $R(t+a) > R(t) \mid a > 0$	Sim	Não
Modelo de Márkov em que há transição partindo do estado de penalidade corresponde a..	Confiabilidade	Disponibilidade
Algoritmo 1º passo, complexidade temporal	$O(n)$	$O(\log(n))$
Algoritmo 1º passo, complexidade espacial	$O(n)$	$O(\log(n))$, versão $O(1)$ não mostrada, mas comentada
Unidade do método dos conjuntos conectores	Elementos (equipamento ou enlace)	Subsistemas onde há elementos redundantes; e elementos que não pertencem a nenhuma disposição com redundância
Algoritmo 2º passo	Não otimizado	Otimizado
Algoritmo 2º passo, o que considera indisponível	Caminho com elemento dormiente	Nenhum subsistema é considerado indisponível. Cada subsistema contabiliza em si os efeitos da dormência de elementos

Quadro 1 - Comparação entre os métodos REASoN e Reasoning

O segundo passo do Reasoning é análogo ao do REASoN, mas ele herda uma diferença da diferença que há entre o primeiro passo desses métodos. O segundo passo aqui é utilizar, no método dos conjuntos conectores ou desconectores, a confiabilidade ou disponibilidade de cada subsistema para calcular a confiabilidade ou disponibilidade referente à totalidade da rede. Uma diferença em relação ao REASoN é que lá o equivalente do subsistema é sempre um único elemento da rede (roteador ou comutador). Outra diferença é que lá o segundo passo “considera como indisponível qualquer **caminho** composto [...] por algum roteador em reserva (ou inativo)” (Amaral, 2012; o destaque é desta dissertação). Essa consideração está ausente no Reasoning. No Reasoning, um elemento em reserva está sempre em um subsistema da rede cuja confiabilidade e disponibilidade tanto depende desse elemento, quanto será considerada pelo método dos conjuntos conectores ou desconectores.

A maior diferença entre o REASoN e o Reasoning é este não permitir, no modelo de confiabilidade, que o sistema saia do estado de Márkov correspondente ao *Delay*. A implicação para a confiabilidade é significativa, pois esta passa a ser equivalente à do caso em que não há redundância. As Figura 4 e Figura 5 ajudam a ilustrar o cenário para o cálculo da disponibilidade, num sistema sem reparo. A Figura 4 mostra três disposições de subsistemas que podem utilizar elementos redundantes em reserva. A Figura 5 mostra os modelos de Márkov para disponibilidade (não para confiabilidade), correspondentes às disposições da Figura 4, no caso em que o modo dormente é aplicado para enlaces e no caso em que ele é aplicado a roteadores ou comutadores.

Figura 4 - Algumas disposições que permitem elementos (enlaces ou dispositivos) redundantes

No subsistema do tipo (a) (Figura 4 (a)), R1 não pode dormir, pois é a única entrada (ou saída) do subsistema. Mas, um dentre R2 e R3 pode dormir, já que ambos são saídas (respectivamente, entradas), sendo desnecessário que ambos estejam ativos simultaneamente. Quando analisando essa disposição do ponto de vista de falhas de dispositivos, deve-se então considerar a possibilidade de haver falha nos três. Do ponto de vista de enlaces, ambos E12 e E13 podem dormir (mas não simultaneamente) ou falhar. No subsistema do tipo (b), do ponto de vista de roteadores, apenas R2 pode dormir, não estando sujeito a falha enquanto dorme, ao passo que R1 e R3 estão sempre sujeitos a falha. Do ponto de vista de enlaces, ou E12 e E23 estão ativos e ambos suscetíveis a falha, de modo que E13 possa dormir, ou E13 está ativo, de modo que os outros enlaces possam dormir (neste caso, R2 também poderia dormir, mas não se lida aqui com modelos onde se imiscuem os efeitos de dispositivo e enlace). No subsistema do tipo (c), nem R1 nem R2 podem dormir, pois ambos são entrada e saída do subsistema. Entretanto, ou E12a ou E12b podem dormir.

Figura 5 - Modelos de Márkov para cálculo de disponibilidade para as disposições da Figura 4

Da Figura 5 vê-se que os modelos que mais se assemelham ao modelo do REASoN para cálculo de confiabilidade (Figura 3 (a)), dum ponto de vista formal, são os modelos (a1) e (c1) para cálculo de disponibilidade (não de confiabilidade). Eles são, respectivamente, os modelos para as disposições (a) e (c) (Figura 4) e ambos consideram que apenas enlaces podem falhar ou dormir. Eles são semelhantes ao modelo da Figura 3 (a) porque também possuem taxa de falhas de apenas um elemento entre S0 e S1 (respectivamente, 1S e *Delay* na Figura 3), de apenas um elemento entre S2 e S3 (respectivamente, 01 e 00), e taxa de despertar de apenas um elemento entre S1 e S2.

As expressões de estado dos modelos (a1) e (c1), para $\lambda_{12} = \lambda_{13} = \lambda$, $\mu_{13} = \mu$, $\lambda_{12a} = \lambda_{12b} = \lambda$ e $\mu_{12b} = \mu$, são:

$$P_0(t + \Delta t) = P_0(t) - \lambda \cdot \Delta t \cdot P_0(t) \quad (\text{Equação 5.4a})$$

$$P_1(t + \Delta t) = P_1(t) - \mu \cdot \Delta t \cdot P_1(t) + \lambda \cdot \Delta t \cdot P_0(t) \quad (\text{Equação 5.4b})$$

$$P_2(t + \Delta t) = P_2(t) - \lambda. \Delta t. P_2(t) + \mu. \Delta t. P_1(t) \quad (\text{Equação 5.4c})$$

$$P_3(t + \Delta t) = P_3(t) + \lambda. \Delta t. P_2(t) \quad (\text{Equação 5.4d})$$

5.3.1 Solução analítica

Será mostrado nesta seção o procedimento para obtenção da solução analítica do modelo (a1) e (c1), os quais são formalmente semelhantes ao modelo-exemplo do REASoN. Manipulando-se algebricamente a Equação 5.4a obtém-se

$$\frac{P_0(t + \Delta t) - P_0(t)}{\Delta t} + \lambda. P_0(t) = 0$$

Tomando-se o limite para $\Delta t \rightarrow 0$, tem-se a seguinte equação ordinária de primeira ordem:

$$\frac{dP_0(t)}{dt} + \lambda. P_0(t) = 0 \quad (\text{Equação 5.5})$$

A solução da Equação 5.5 é da forma $A_1 e^{-\lambda t}$, com A_1 constante (Zill e Cullen, 2001). Da condição inicial $P_0(0) = 1 \leftrightarrow A_1 e^{-\lambda \cdot 0} = 1$, donde $A_1 = 1$, e a solução analítica para a probabilidade do sistema estar no estado lógico S_0 é, então

$$P_0(t) = e^{-\lambda t} \quad (\text{Equação 5.6})$$

O procedimento para resolução quanto ao estado S_1 é análogo. Manipula-se algebricamente a Equação 5.4b e toma-se o limite para $\Delta t \rightarrow 0$, de forma a se obter

$$\frac{dP_1(t)}{dt} = \lambda. P_0(t) - \mu. P_1(t) \quad (\text{Equação 5.7})$$

Levando-se em conta a expressão de $P_0(t)$ (Equação 5.6), a solução de $P_1(t)$ é da forma $P_1(t) = B_1 e^{-\lambda t} + B_2 e^{-\mu t}$. Substituindo tal forma na Equação 5.7, vem

$$-\lambda B_1 e^{-\lambda t} - \mu B_2 e^{-\mu t} = \lambda e^{-\lambda t} - \mu B_1 e^{-\lambda t} - \mu B_2 e^{-\mu t}$$

Muitos termos se cancelam e, resolvendo para B_1 , tem-se $B_1 = \frac{\lambda}{\mu - \lambda}$. Das condições iniciais, $P_1(0) = 0 \leftrightarrow B_2 = -B_1$ e a solução analítica para $P_1(t)$ é dada pela Equação 5.8.

$$P_1(t) = \frac{\lambda}{\mu - \lambda} (e^{-\lambda t} - e^{-\mu t}) \quad (\text{Equação 5.8})$$

O procedimento para obtenção das soluções analíticas para $P_2(t)$ e $P_3(t)$ é semelhante, mas mais extenso, de forma que não será mostrado aqui. O material de resolução de equações diferenciais por meio do uso de fator de integração apresentado por Zill e Cullen (2001) é de uso especial nessa tarefa.

5.3.2 Algoritmo iterativo

O Algoritmo 5.1 mostra um modo básico de se efetuar o procedimento iterativo. Sendo o método proposto voltado para ser usado por soluções sustentáveis e também para a avaliação de tais soluções, é cogente que ele também seja, na medida do possível, otimizado. Antes de se apresentar o cerne de uma versão mais eficiente do Algoritmo 5.1, apresentam-se alguns ajustes menores. Esses ajustes foram aplicados na obtenção do Algoritmo 5.3. São eles:

1. Mudança nas expressões das probabilidades relacionadas a cada estado, segundo as alterações no modelo de Márkov. Note-se que, apesar das expressões mostradas no Algoritmo 5.3 serem semelhantes às do Algoritmo 5.1, devido às diferenças dos métodos correspondentes, elas calculam grandezas diferentes (respectivamente, disponibilidade dos modelos (a1) e (c1) da Figura 5 e confiabilidade do modelo da Figura 3A).
2. A inicialização concernente às condições de contorno são realizadas fora do laço, ainda que o intervalo $[0, t_{\text{fim}}-1]$ da variável de controle tenha sido mantido. Esta mudança também acarreta ganho de desempenho na

execução do laço, na grande maioria dos processadores existentes e comuns¹².

3. O cálculo das grandezas confiabilidade e disponibilidade é realizado apenas uma vez, ao final do processo.

O Algoritmo 5.3 assume que t_{fim} é múltiplo de Δt , ou seja, $t_{fim} = n\Delta t$, $n \in \mathbb{N}^+$. Esse algoritmo mostra o cálculo para a disponibilidade dos modelos (a1) e (c1). De fato, uma simulação para $\Delta t = 1$, $\mu = 0,8$, $\lambda = 0,5$ e $t_{fim} = 3$ mostrará que esse modelo é mais adequado para cálculo de disponibilidade do que para cálculo de confiabilidade: no instante 0, $P_0 = 1$ e $P_1 = P_2 = P_3 = 0$. No instante 1, $P_0 = 0,5$, $P_1 = 0,5$ e $P_2 = P_3 = 0$, de forma que $P_0 + P_2 = 0,5$, ou seja, a probabilidade do sistema estar operando adequadamente é 0,5. No instante 2, $P_0 = 0,5 - 0,5*0,5 = 0,25$, $P_1 = 0,5 - 0,8*0,5 + 0,5*0,5 = 0,35$, $P_2 = 0,4$ e $P_3 = 0$. Assim, no instante 2 temos $P_0(2) + P_2(2) = 0,65 > P_0(1) + P_2(1) = 0,5$. Deste modo, a soma $P_0 + P_2$ não pode corresponder a uma conta de confiabilidade, pois, neste caso, teríamos que a probabilidade do sistema funcionar satisfatoriamente até o instante 2 seria maior que a probabilidade do sistema funcionar satisfatoriamente até o instante 1. Logo, os modelos da Figura 5 e seu correspondente (Figura 3A) não podem corresponder a confiabilidade. Podem, entretanto, corresponder ao modelo de disponibilidade. E, de fato, correspondem.

Algoritmo 5.3: Versão ajustada do Algoritmo 5.1

$P_0(0) = 1$; $P_1(0) = P_2(0) = P_3(0) = 0$

Laço de $t = 0$ a $t = t_{fim} - \Delta t$

$$P_0(t + \Delta t) = P_0(t) - \lambda \cdot \Delta t \cdot P_0(t)$$

$$P_1(t + \Delta t) = P_1(t) - \mu \cdot \Delta t \cdot P_1(t) + \lambda \cdot \Delta t \cdot P_0(t)$$

$$P_2(t + \Delta t) = P_2(t) - \lambda \cdot \Delta t \cdot P_2(t) + \mu \cdot \Delta t \cdot P_1(t)$$

$$P_3(t + \Delta t) = P_3(t) + \lambda \cdot \Delta t \cdot P_2(t)$$

$$t = t + \Delta t$$

Fim Laço

$A_{fim} = A(t) = P_0(t) + P_2(t)$

¹² Ainda que a otimização expressa nesse item seja trivial, dependendo da implementação que o programador escolher para as variáveis, a otimização por parte do compilador pode não ocorrer, mesmo para o nível -o3 do GCC (Costa, 2012b).

Seja $n = t_{fim}/\Delta t$ o número de iterações executadas no Algoritmo 5.3. Quando $\Delta t = 1$ segundo, o cálculo para dias ou meses pode ficar muito demorado. Quando se deseja calcular a disponibilidade para um período de um ano, faz-se então necessário um modo mais rápido de se efetuar o cálculo iterativo. Tal modo parte da percepção de que as contas efetuadas a cada iteração do Algoritmo 5.3 correspondem a uma multiplicação matricial. Assim, seja o vetor $P(t) = [P_0(t), P_1(t), P_2(t), P_3(t)]$. Então, a cada iteração, o Algoritmo 5.3 calcula $P(t+\Delta t) = A.P(t)$, com A dado pela Equação 5.9.

$$A = \begin{pmatrix} 1 - \lambda.\Delta t & 0 & 0 & 0 \\ \lambda.\Delta t & 1 - \mu.\Delta t & 0 & 0 \\ 0 & \mu.\Delta t & 1 - \lambda.\Delta t & 0 \\ 0 & 0 & \lambda.\Delta t & 1 \end{pmatrix} \quad (\text{Equação 5.9})$$

Deste modo, para o caso de $\Delta t = 1$, então $P(1) = A.P(0)$, $P(2) = A.P(1) = A^2.P(0)$, até $P(n) = A^n.P(0)$. Ora, essa multiplicação matricial não precisa ser realizada n vezes. Basta notar que $A^n = A^{n/2} \times A^{n/2}$. Decompondo a exponenciação desta forma, são necessárias $n/2$ iterações para o cálculo de $A^{n/2}$ e mais outra para o cálculo de $A^{n/2} \times A^{n/2}$. Prosseguindo com o raciocínio da decomposição, uma próxima melhoria seria realizarem-se $n/4$ iterações para o cálculo de $A^{n/4}$, mais uma para o cálculo de $A^{n/2} = A^{n/4} \times A^{n/4}$ e a última para o cálculo de A^n . Esse raciocínio diminui substancialmente a complexidade original de $O(n)$ operações matriciais, com um custo adicional de memória nulo. No melhor caso, $n = 2^m$ e o número de operações matriciais necessárias é $\log(n) = m$. Nesse caso, a complexidade espacial é $O(1)$, pois só se necessita armazenar o valor da variável auxiliar, como mostra a Listagem 1:

Listagem 1 – Execução eficiente correspondente ao Algoritmo 5.3,
para $\Delta t = 1$ e $t_{fim} = 2^m$

Operação 0 – $Aux \leftarrow A$
 Operação 1 – $Aux \leftarrow Aux.Aux$ // $Aux = A^2$
 Operação 2 – $Aux \leftarrow Aux.Aux$ // $Aux = A^4$
 Operação 3 – $Aux \leftarrow Aux.Aux$ // $Aux = A^8$
 ...
 Operação m – $Aux = Aux.Aux$ // $Aux = A^{2^m}$

No pior caso, $n = 2^m + 2^{m-1} + 2^{m-2} + \dots + 2^0$ (por exemplo, $n = 15 = 8 + 4 + 2 + 1$). Nesse caso, são necessárias $2m$ operações matriciais e a complexidade espacial é m , como indica a Listagem 2. Essa complexidade temporal de $2m$ do pior caso pode ser reduzida a m , ao mesmo tempo em que complexidade espacial é reduzida a 2. Assim, ao invés de se usarem algoritmos distintos para o melhor e pior casos, pode-se obter um algoritmo único de desempenho semelhante para ambos os casos. Ainda, para o melhor caso esse algoritmo tem o custo espacial adicional de apenas 1 (uma matriz) em relação ao cálculo mostrado na Listagem 1. Tal é o Algoritmo 5.4.

Listagem 2 – Idem à Listagem 1, mas para $t_{fim} = 2^m + 2^{m-1} + \dots + 2^0$

Operação 0 – $Aux0 \leftarrow A$
 Operação 1 – $Aux1 \leftarrow Aux0.Aux0$ // $Aux1 = A^2$
 Operação 2 – $Aux2 \leftarrow Aux1.Aux1$ // $Aux2 = A^4$
 Operação 3 – $Aux3 \leftarrow Aux2.Aux2$ // $Aux3 = A^8$
 ...
 Operação m – $Auxm \leftarrow Auxm-1.Auxm-1$ // $Auxm = A^{2^m}$
 Operação m+1 – $Auxm \leftarrow Auxm.Aux0$
 Operação m+2 – $Auxm \leftarrow Auxm.Aux1$
 ...
 Operação m+m – $Auxm \leftarrow Auxm.Aux_{m-1}$

Algoritmo 5.4: Cálculo eficiente correspondente ao Algoritmo 5.3

Entrada

Δt passo de iteração, assumido divisor de *tempo final*
 λ taxa de falhas
**tempo final* instante final, para o qual se deseja a disponibilidade do sistema
 N $\text{tempo final} / \Delta t$
 $P(0)$ vetor da probabilidade inicial do sistema estar em cada estado

* grandezas de mesma dimensão temporal que $1/\lambda$ (por exemplo, segundos)

** no modelo em pauta, $P(0) = [P_0(0), P_1(0), P_2(0), P_3(0)] = [1, 0, 0, 0]$

Saída

D a disponibilidade do sistema no instante desejado

```

00  $k \leftarrow N, A \leftarrow$  Equação 5.9, para  $\Delta t$  e  $\lambda$ 
01  $A_{\text{fim}} \leftarrow I$  //  $A_{\text{fim}}$  recebe a matriz identidade
02 Se  $k = 0$  Então
03   Vá para a Linha 13
04 Fim Se
05  $Aux \leftarrow A$  //  $Aux$  fica definido pela Equação 5.9
06 Faça
07   Se  $((k \& 1) == 1)$  Então
08      $A_{\text{fim}} \leftarrow Aux \times A_{\text{fim}}$ 
09   Fim Se
10    $k \leftarrow (k \gg 1)$  //  $N$  é deslocado um bit à esquerda
11    $Aux \leftarrow Aux \times Aux$ 
12 Enquanto  $k > 0$ 
13    $P_{\text{final}} = A_{\text{fim}} \times P(0)$ 
14 Retorna  $D = P_{\text{final}}(0) + P_{\text{final}}(2)$ 
  
```

O Algoritmo 5.4 corresponde ao primeiro passo do método Reasoning. Esse algoritmo é uma versão otimizada do Algoritmo 5.3, que por sua vez é uma extensão ao Algoritmo 5.1, do método REASoN. O Algoritmo 5.4 utiliza uma única variável para ir calculando algumas potências de A (Linha 11). Essa matriz dada

pela Equação 3.9. As potências calculadas de A são da forma A^n , com $n = 2^x$, tal que $0 \leq 2^x \leq N$, sendo N a razão *tempo final* / Δt . N corresponde ao número de iterações que seriam gastas pelo Algoritmo 5.3. Ao passo que essas potências de A são calculadas, uma segunda variável auxiliar, A_{fim} , vai sendo atualizada mediante multiplicação com alguma potência de A (Linha 08). Os momentos de atualização são determinados pela operação lógica *E bit a bit*, realizada entre a variável de controle k e 1 (Linha 07). Ou seja, a Linha 07 checa se o bit menos significativo de k está ativo. k é inicialmente carregada com N (Linha 00) e, a cada iteração, k é deslocada um bit à direita (Linha 10). As Linhas 00 – 05 correspondem à inicialização da lógica principal do algoritmo, que é o laço encerrado pelas Linhas 06 e 12. Quando $N > 0$, o Algoritmo 5.4 sempre acaba executando uma operação matricial em vão antes de sair do laço (Linha 11), pois o novo valor de Aux não é utilizado.

5.4 Considerações do capítulo

Este capítulo descreveu o método Reasoning. Este método deve ser usado para calcular a disponibilidade ou confiabilidade da rede nas situações em que há dispositivos ou enlaces que entram no modo dormiente ou quejando. O método pode ser usado para determinar se uma mudança para uma nova configuração é aceitável ou para avaliar a nova configuração já aplicada. Desta forma, o método ajuda a determinar se os acordos de nível de serviço continuam satisfeitos. O Reasoning produz um resultado mais acurado que o REASoN no cálculo da confiabilidade, por não só considerar o tempo gasto em acordar como um período de indisponibilidade, mas também refletir essa consideração nos modelos de Márkov. Do ponto de vista de confiabilidade, o tratamento dado pelo Reasoning mostra que a situação é semelhante ao dispositivo dormiente não existir. Várias técnicas de eficiência energética (por exemplo, uso de ACPI e coalescimento síncrono no comutador) causam um cenário em que elementos de rede entram no

modo dormente, e vários algoritmos de roteamento voltados para eficiência energética (por exemplo, ElasticTree e GreeTE) roteiam de modo a ocasionar a possibilidade de elementos de rede dormirem. Assim, o método Reasoning lida com um cenário que deve ser comum em redes voltadas a eficiência energética.

Apresentou-se a quais disposições de elementos correspondem diferentes modelagens de Márkov. O método Reasoning utiliza um algoritmo iterativo mais eficiente que seu predecessor no cálculo da confiabilidade ou disponibilidade. O algoritmo iterativo pode ser utilizado quando as soluções às equações diferenciais correspondentes às probabilidades de estado do modelo de Márkov não estiverem disponíveis.

6 Arquitetura e implementação de um sistema para eficiência energética

O Capítulo 3 descreveu técnicas proeminentes para a implementação de redes de computadores mais sustentáveis. Se por um lado essas técnicas aumentam a eficiência energética da rede, por outro elas alteram o uso comum dos equipamentos, acarretando um compromisso entre ganhos de energia e degradação na qualidade de serviço. Também mostrou-se de que modo os equipamentos podem vir a ter a vida útil encurtada. Há também técnicas verdes que não necessariamente afetam tanto os equipamentos, mas que tão só buscam agregar o tráfego ou enviar cada fluxo de dados pelo caminho energeticamente mais eficiente. Assim, é próprio do universo das técnicas verdes que o sistema de gerenciamento de rede facilite a interação entre os diferentes compromissos e a aplicação das próprias decisões verdes. Tal sistema deve lidar com informações de diferentes abstrações, como objetivos econômicos, restrições de conectividade e acordos de nível de serviço.

Este capítulo descreve a arquitetura e detalhes da implementação do SustNMS (Carlos, 2012), bem como descreve sua dinâmica. O SustNMS é um sistema de gerenciamento de rede orientado por políticas de sustentabilidade. Os testes das capacidades de técnicas de eficiência energética do SustNMS foram feitos em um ambiente de experimentação (Januário, 2013). Apresenta-se neste capítulo a arquitetura do referido sistema e discorre-se sobre o método de sua implementação.

6.1 Visão geral do sistema SustNMS

A arquitetura do SustNMS, acrônimo para *Sustainability-Oriented Network Management System* (Sistema de Gerenciamento de Rede Voltado a Sustentabilidade), é uma arquitetura de um sistema de gerenciamento de redes voltado a sustentabilidade. Ela guia a construção de um sistema que avalia e controla a operação da rede com base em indicadores de sustentabilidade, disponibilidade e desempenho expressos a partir de políticas de gerenciamento do sistema. No caso de sustentabilidade, são considerados fundamentalmente indicadores de eficiência energética e previstos indicadores de ciclo de vida. No caso de qualidade de serviço, relacionada a desempenho e disponibilidade, são considerados indicadores de taxa de transmissão efetiva, atraso, taxa de perda de pacotes, entre outros.

Esse sistema de gerenciamento voltado a sustentabilidade pressupõe a avaliação de eficiência energética em redes heterogêneas, incluindo redes onde haja dispositivos incapazes de prover informações relevantes, como o consumo de energia.

Os elementos da rede (roteadores e comutadores) são gerenciados pelo sistema SustNMS de modo que se maximize a utilização sustentável dos recursos. Os critérios para tal maximização são descritos por meio de políticas. As políticas são expressas em alto-nível pelo administrador da rede, de modo que, para serem implementadas em cada equipamento onde atuem, precisem ser intermediadas por um ponto de aplicação de política. Os índices de sustentabilidade são computados com base em indicadores de consumo coletados dos elementos da rede em tempo real.

A arquitetura discutida contempla também um módulo para dinamicamente calcular a disponibilidade da rede quando diferentes nós e enlaces são desativados ou postos em diferentes modos dormentes. Quando em modo dormente, o equipamento consome menos energia, acarretando uma rede mais sustentável. Entretanto, nesse modo o equipamento não perfaz todas suas funcionalidades operacionais, o que diminui a confiabilidade e disponibilidade da

rede. Justamente por isso, a arquitetura discutida conta com esse módulo para permitir a avaliação de compromissos entre disponibilidade e sustentabilidade.

6.2 Arquitetura

A arquitetura do SustNMS está mostrada na Figura 6. O sistema requer que os dispositivos tenham certas funcionalidades habilitadas no ambiente de experimentação (mais detalhes na Seção 7.1). O SustNMS estende a arquitetura de gerenciamento baseado em política definida pela IETF ao incluir-lhe três módulos: o **Repositório de Modelos**, o **Monitor de Qualidade de Serviço** e o **Monitor de Sustentabilidade**.

O primeiro módulo incluído na arquitetura do SustNMS (Figura 6), o **Repositório de Modelos**, contém dois submódulos:

- a) um para armazenamento de modelos de consumo energético, os quais são parâmetros estáticos que definem perfis de consumo de energia;
- b) o outro para armazenamento de modelos de disponibilidade. Modelos de disponibilidade abrangem tanto informações de taxas de falha e de reparo, quanto o modo como as taxas dos diferentes dispositivos da rede se inter-relacionam.

O segundo módulo incluído, o **Monitor de Qualidade de Serviço**, também contém dois submódulos:

- a) um para dinamicamente avaliar a disponibilidade da rede a cada vez que uma nova configuração é calculada pelo arcabouço de gerenciamento de política;
- b) outro para dinamicamente coletar indicadores de desempenho de cada nó da rede.

Figura 6 - Arquitetura do SustNMS (Costa, 2012a)

O terceiro módulo incluído, o **Monitor de Sustentabilidade**, é composto por um único submódulo, o Avaliador de Eficiência Energética (AEE). Essa divisão permite que se incluam outros módulos de avaliação. Por exemplo, a arquitetura pode ser expandido de modo a incluir um módulo que avalie o tempo de vida dos equipamentos, o que poderia ser realizado por meio do método descrito no Capítulo 4. O AEE avalia o consumo de energia instantâneo de cada nó, assim permitindo uma avaliação da eficiência energética de toda a rede.

O **Atualizador de Dispositivo** (AD) é o módulo responsável por coletar dados e por aplicar mudanças a cada dispositivo da rede (por exemplo, de estado de consumo de energia). Para tanto, o AD comunica-se com cada instância de **Ponto de Aplicação de Política** (PAP). Os PAPs são responsáveis por executar as ações necessárias para que a rede contemple as políticas de gerenciamento definidas pelo gerente da rede. Os PAPs podem realizar validação específica de cada dispositivo, se necessário.

Tendo-se visto em que aspectos a arquitetura do SustNMS difere-se da arquitetura da IETF, discorre-se agora sobre esta. A arquitetura previamente definida pela IETF, ou seja, a arquitetura subjacente às adições comentadas acima, é composta pelo Arcabouço de Gerenciamento de Política, por sua vez composto por três submódulos (Costa, 2012a):

- a) Ferramenta de Gerenciamento de Política, uma ferramenta para edição e validação;
- b) Ponto de Decisão de Política, um módulo que detecta quando uma regra de política deve ser aplicada e garante que ela o seja; e
- c) Repositório de Política, correspondente a um banco de dados.

6.3 Descrição dos submódulos da arquitetura do SustNMS

O objetivo do SustNMS é gerenciar os roteadores e comutadores de uma rede. Cada um destes está representado na arquitetura da Figura 6 pelo módulo Comutador/Roteador. Assume-se que esse módulo contém ao menos uma funcionalidade de economia de energia. Um exemplo de tal funcionalidade é o suporte ao estado dormiente, quando no qual um elemento de rede correspondente não está realizando tarefas relevantes, como o encaminhamento de pacotes de dados, mas ainda se encontra minimamente energizado. O elemento de rede deve estar minimamente energizado, pois é mister que ele possa acordar do estado dormiente após um estímulo enviado através da rede de controle.

Para o módulo **Comutador/Roteador**, a arquitetura prevê os seguintes submódulos:

- a) Dados de Tráfego, responsável pela coleta de dados de tráfego de cada interface do dispositivo;
- b) Protocolo de Gerenciamento, representando um protocolo, como o SNMP, capaz de juntar os dados de tráfego coletados e atualizar uma base de dados relativos a gerenciamento, como a MIB (*Management Information Base*, Base de Informações de Gerenciamento), do próprio SNMP. É também assumido que esse submódulo tem controle sobre o estado energético, velocidade e banda das interfaces.

No **Ponto de Decisão** reside o controle do sistema. O **Ponto de Decisão** atua consoante uma política de gerenciamento e compromissos aceitáveis. Esse módulo utiliza um grafo de rede para concentrar toda a informação e índices calculados da rede. Nesse grafo estão representados cada comutador e roteador gerenciado, sendo os enlaces e caminhos entre os nós representados pelos arcos do grafo. As propriedades dos nós e arcos do grafo são representadas por tais variáveis:

- NO.ID: identificador único do nó;

- NO.CLASSE: identificador da classe do nó;
- NO.NECR: taxa de consumo energético de rede (*Network Energy Consumption Rate*)(Manral, 2010), medido em miliwatts por Mbps;
- NO.NEPI: índice de proporcionalidade de energia de rede (*Network Energy Proportionality Index*)(Manral, 2010);
- NO.ELTPI: índice de proporcionalidade do tempo de vida esperado (*Expected Lifetime Proportionality Index*);
- NO.PKR: taxa de perda de pacote (*Node Packet Loss Rate*) do caminho mínimo do equipamento;
- NO.DELAY: atraso entre um pedido de um dispositivo e a resposta da rede de transporte;
- NO.AVAIL: índice de disponibilidade do nó;
- ARCO.BANDWIDTH: largura de banda do enlace; e
- ARCO.ESTADO: estado do arco (dormente, ativo, com defeito).

As informações do grafo da rede devem ser atualizadas pelo **Ponto de Decisão**. Este módulo aciona o **Monitor de Sustentabilidade**, o **Monitor de Disponibilidade** e o **Atualizador de Dispositivo**. A taxa de amostragem determina a velocidade de atualização dos dados. Essa taxa é escolhida pelo usuário do sistema de gerenciamento.

Na atualização, primeiramente, o **Ponto de Decisão** aciona o **Atualizador de Dispositivo**, o qual é o módulo responsável por coletar dados e aplicar mudanças de configuração em cada comutador ou roteador da rede. Para tanto, o **Atualizador de Dispositivo** deve comunicar-se a cada instância de Protocolo de Gerenciamento. O **Atualizador de Dispositivo** modifica os valores das variáveis associadas a cada arco do grafo, de sorte que reflitam a carga e o estado

energético de cada enlace. A partir da consulta do Protocolo de Gerenciamento, o módulo **Atualizador de Dispositivo** também coleta dados sobre o desempenho do nó. Tais dados incluem a taxa de perda de pacotes e o atraso na comunicação entre o **Atualizador de Dispositivo** e os dispositivos a serem atualizados.

O **Ponto de Decisão** aciona o **Monitor de Sustentabilidade** para que este avalie as informações do grafo de rede. A avaliação gera outros índices, que são também marcados no grafo. O **Monitor de Sustentabilidade** percorre o grafo e, para cada nó, ele aciona o submódulo **Avaliador de Eficiência Energética**. Dada a taxa de amostragem, o **Avaliador de Eficiência Energética** pode calcular a taxa de consumo de energia.

O próximo passo é verificar se há no **Repositório de Modelo de Energia** algum modelo para o dispositivo em questão. Não havendo um modelo definido, o sistema utiliza um modelo de consumo pessimista. O próximo passo é processar a **Avaliação de Eficiência Energética**, processo no qual os modelos de energia e as demais informações do grafo são utilizados. O tráfego de cada nó é calculado pelo tráfego associado a cada enlace relacionado às portas do nó.

Na sequência, o **Ponto de Decisão** aciona o módulo **Avaliador de Disponibilidade**. Este módulo percorre o grafo de informações da rede para consultar os modelos de disponibilidade e calcular a disponibilidade, primeiramente de cada nó, posteriormente da rede toda. O cálculo de disponibilidade adotado pelo SustNMS no contexto desta dissertação é o método de dois passos REASoN. Esta dissertação recomenda o uso do método, também de dois passos, Reasoning (sobre esses métodos, veja-se o Capítulo 5). As informações contidas no Modelo de Disponibilidade são:

- AV.ENLACE_MTBF: tempo médio até a falha de um enlace;
- AV. ENLACE _MTTR: tempo médio de reparo de um enlace;
- AV. ENLACE _WAKUP_TIME: demora em despertar do enlace;

- AV.NOS_MTBF: tempo médio até a falha de um nó;
- AV. NOS _MTTR: tempo médio de reparo de um nó; e
- AV. NOS _WAKUP_TIME: demora em despertar do nó.

O **Ponto de Decisão** então percorre o grafo atualizado pelos **Monitores de Sustentabilidade e Disponibilidade**, assinalando pesos que refletem decisões que poderiam ser tomadas. O processo todo é repetido a cada amostragem realizada pelo SustNMS. A acurácia das medidas é definida pela taxa de amostragem. A taxa adequada acarreta mudanças na própria eficiência energética (é inevitável aqui uma comparação com a física quântica), já que um aumento da taxa de amostragem refletiria num acréscimo das mensagens trocadas pelas operações de gerenciamento, ulteriormente levando a um acréscimo no consumo de energia.

6.4 Detalhes da implementação

Para avaliação do sistema proposto, a arquitetura apresentada na Seção 6.2 foi implementada em Python. O SustNMS é executado sobre o plano de dados de uma rede *Multi Protocol Label Switching* (MPLS). Esta é uma tecnologia amplamente utilizada por operadoras de redes comerciais (Kempf, 2011). Os possíveis caminhos utilizáveis pelo MPLS são definidos de antemão, de sorte que não é necessário calcular rotas durante a operação da rede. Consoante descrito por Dongmei e Guangzhi (2008), a utilização de caminhos de reserva previamente estabelecidos permite rerroteamento rápido e baixo sobrecusto no plano de controle.

O sistema é inicializado com o carregamento do **Ponto de Decisão de Política** (PDP), o qual é conectado a um único **Atualizador de Dispositivo** (AD) e a um único **Avaliador de Eficiência Energética** (AEE). As ações regulares do PDP são

requerer periodicamente que o AD inicie um processo de coleta de informações dos equipamentos e requerer que o AEE avalie a situação da rede, após a representação interna da informação de todos os dispositivos houver sido coletada. Informações estáticas como perfil de consumo e modelo de disponibilidade de cada dispositivo são armazenadas no Repositório de Modelos. Informações que mudam com o tempo, como a carga ou o modo de consumo de algum dispositivo, por variarem dinamicamente de acordo com a demanda dos clientes da rede, são acessadas pelo AD via SNMP.

Após a etapa de requisição de informações (ou amostragem), o **Monitor de Qualidade de Serviço** avalia a disponibilidade e o desempenho da rede, permitindo ao **Ponto de Decisão de Política** checar se as restrições de política permanecem satisfeitas após a aplicação do Algoritmo 6.1, abaixo. Esse algoritmo é aplicado na definição de quais túneis MPLS e, conseqüentemente, de quais dispositivos devem estar ativos para que a rede supra a demanda de tráfego corrente. O **Repositório de Política** e os **Repositórios de Modelos** de disponibilidade e de consumo energético são armazenados num banco de dados MySQL.

O Algoritmo 6.1 é utilizado para garantir que todos os fluxos serão atendidos pela solução de roteamento adotada. No caso, um fluxo é determinado pelo roteador de entrada na rede, pelo roteador de saída e pela carga a ser conduzida. Então, por exemplo, se a carga no fluxo 2 for nula, S_2 não requer que nenhum caminho específico esteja ativo, sendo, portanto, nulo. O Algoritmo 6.1 lida apenas com dois pares distintos de roteadores de entrada e de saída. Esse algoritmo também garante a conectividade da rede. Sua função *consumo(x,y)* representa uma consulta ao modelo de consumo do dispositivo y . Nota-se que o próprio modo de definição das soluções candidatas já embute em si uma política, a saber, a política de que uma solução candidata deve manter a conectividade dos clientes (os quais podem gerar os fluxos 1 e 2).

Algoritmo 6.1: Mecanismo de escolha de quais túneis MPLS
predefinidos o SustNMS usará

Variáveis

c um conjunto de caminhos, denotando uma solução candidata
 R um roteador
 l_R carga em R
 P_c consumo da solução candidata c
 $\text{consumo}(x,y)$ retorna o consumo de y sob carga x

Saída o conjunto de caminhos que devem estar ativos

$S_0 \leftarrow$ conjunto de conjuntos de roteadores que garantem a conectividade da rede
 $S_1 \leftarrow$ conjunto de caminhos que garantem o fluxo 1
 $S_2 \leftarrow$ conjunto de caminhos que garantem o fluxo 2
 $CS \leftarrow \{C \in A \cup B \cup C \mid A \in S_0, B \in S_1, C \in S_2\}$
 $P_{\min} = \infty$
 $C_{\min} = 0$
Para cada c em CS
 Se c satisfaz política **então**
 $P_c = 0$
 Para cada R em c
 $l_R \leftarrow$ carga imposta em R pela solução
 $P_c \leftarrow P_c + \text{consumo}(l_R, R)$
 Fim para cada R
 Se $P_c < P_{\min}$ **então**
 $P_{\min} \leftarrow P_c$
 $C_{\min} \leftarrow C$
 Fim Se
 Fim Se
Fim para cada c
Retorna $C_{\min}$

6.5 Considerações do capítulo

Neste capítulo foram apresentadas a arquitetura e a implementação de um sistema de gerenciamento de rede orientado por política de sustentabilidade, o SustNMS. Por ser baseado em política, o sistema fornece a gerentes de redes um controle em alto nível, o qual é preferível a controles feitos por interação direta de gerentes com equipamentos de rede. O sistema apresentado permite que se estabeleçam políticas cientes da disponibilidade da rede, mas não incorpora um método para análise do tempo de vida. Como tal sistema faz uso de funcionalidades não disponíveis nos dispositivos comerciais atuais, as quais, entretanto, são desejáveis e necessárias, o sistema demanda um ambiente de experimentação com certas peculiaridades. O método de implementação do ambiente de experimentação e a avaliação do protótipo SustNMS são os temas dos capítulos subsequentes.

7 Ambiente de experimentação para soluções sustentáveis

O presente capítulo apresenta como se estenderem abordagens existentes para realização de experimentos baseados em ambientes de experimentação e voltados a redes atentas ao consumo energético. A extensão consiste em se adicionar, por meio de ferramentas de gerenciamento de rede, o suporte à avaliação de custos e benefícios relacionados à coordenação de técnicas de otimização de consumo de energia. Assim, podem-se avaliar potenciais compromissos entre economizar energia e degradar desempenho em ambientes de rede. O objetivo é definir um ambiente de experimentação que permita a avaliação de técnicas verdes, atentas ao consumo energético, do ponto de vista de gerenciamento. Dessa forma, o ambiente é de especial utilidade em análises que envolvam algoritmos e *hardwares* do futuro, aplicados à infraestrutura de redes de comunicação, posto que técnicas verdes são incipientes. Requisitos para o ambiente são levantados na Seção 7.1, ao passo que a Seção 7.2 apresenta como esses requisitos foram implementados na preparação do ambiente de experimentação do SustNMS. Considerações são apresentadas em 7.3.

7.1 Requisitos para o ambiente de experimentação

Um ambiente de experimentação para avaliação de funcionalidades de redes energeticamente eficientes requer que os roteadores sejam munidos de diferentes estados de consumo de energia e possam fornecer dados sobre seu consumo corrente. Para permitir tal medição em tempo real, modelos de energia podem servir para parametrização do consumo de energia. Um perfil de consumo adequado deve produzir um relatório de consumo acurado quando dados acerca do modo de consumo e do tráfego que atravessa o equipamento são providos. Utilizam-se neste trabalho perfis lineares, determinados pelo consumo de quando

o equipamento não trata nenhuma carga (representando, assim, um consumo fixo) e por um fator de escala relacionado ao aumento de carga (representando, assim, uma variação de consumo, a qual depende do tráfego). Os perfis foram obtidos da literatura; mais detalhes se encontram na Seção 8.4.

Perfis de consumo podem ser usados na emulação de funcionalidades desejadas. Em tais casos, roteadores implementados em *software* são comumente empregados. Uma vez que protocolos padrão de roteamento não necessariamente definem o caminho ótimo acerca de eficiência energética e restrições de QoS, deseja-se um modo de se garantir o uso de tais caminhos. A Figura 7 resume os requisitos para o ambiente de experimentação de técnicas verdes.

Figura 7 - Requisitos para ambiente de experimentação voltado à avaliação de funcionalidades de sistemas de gerenciamento energeticamente eficiente

Os principais requisitos concernentes a um ambiente de experimentação para avaliação de sistemas de gerenciamento energeticamente eficiente são:

- a) Máquinas virtuais: Usar máquinas virtuais diminui a quantidade de equipamentos físicos necessários, tornando os experimentos mais escaláveis, enquanto ainda mantendo um nível de concretude quanto às operações que ocorreriam nos processadores reais equivalentes.
- b) Sistema operacional alterável: Para que seja possível o cumprimento de outros requisitos (por exemplo, de controle de vazão e de rotas forçadas), é

usualmente necessário que se altere o sistema operacional das máquinas virtuais, de modo a se poderem implementar funcionalidades de baixo nível (como protocolos de rede das camadas 2 e 3).

c) Suporte a Técnicas de Eficiência Energética:

- I. Suporte a estado dormiente: As técnicas de eficiência energética alteram a frequência de operação dos enlaces e processadores (cf. Seção 3.2), sendo comum também pô-los em modo dormiente. Um ambiente de experimentação deve prover tal estado.
- II. Controle do tempo de despertar: Como o tempo de despertar é não negligenciável, a emulação das soluções verdes é mais realista se emular demora dos equipamentos para despertar.

d) Rotas Forçadas: Além das soluções que agem localmente nos dispositivos, há as que gerenciam a rede como um todo, tipicamente atuando no roteamento, daí a necessidade do ambiente de experimentação possibilitar que rotas sejam forçadas, não apenas definidas por algum protocolo usual (por exemplo, OSPF).

e) Mecanismos de Aplicação de Decisão: As decisões de roteamento, ajuste de frequência ou mudança de estado energético devem ser realizadas nos dispositivos (no caso, virtuais), de modo que é necessário um mecanismo para aplicação das decisões.

f) Provisão de Informação sobre Pacotes Perdidos: Algumas soluções (por exemplo, o SustNMS) utilizam no mecanismo de decisão a informação sobre perda de pacotes. Um ambiente de experimentação deve possibilitar que o sistema de gerenciamento tenha acesso a essa informação.

g) Provisão de Informação sobre Consumo Energético: Permite que sejam avaliados compromissos entre a economia de energia e degradação de

qualidade de serviço, valiosos em um gerenciamento sustentável guiado por política.

7.2 Implementação do ambiente de experimentação

O ambiente de experimentação para avaliação do SustNMS é baseado em roteadores implementados em software, com sistema operacional GNU/Linux (distribuição Debian). O uso de roteadores baseados em GNU/Linux está crescendo (Chabarek e Barford, 2011). Por isso, o ambiente de testes descrito é aplicável a uma ampla gama de equipamentos. Para emular diferentes roteadores, a rede é construída com máquinas virtuais, usando VMWare vSphere (ESXi). As máquinas virtuais foram interconectadas com VMWare vSwitch (encaminhamento na Camada 2, com rotulação VLAN – *Virtual Local Area Network*). Desabilitar os controladores das interfaces de rede (NICs – *Network Interface Controllers*) concernentes ao plano de dados emula o estado dormente de energia. NICs de plano de controle jamais são desabilitadas, de modo que se mantenha a presença de rede do correspondente do equipamento. Além de se desligarem NICs de plano de dados, o sistema utiliza perfis de consumo para modelar o consumo energético do equipamento quando este se encontra em modo dormente. Neste modo o equipamento não lida com carga, mas pode acordar passando para o modo ativo. No caso do sistema testado, os valores de tempos de troca de estado são também armazenados internamente, no **Modelo de Disponibilidade** do SustNMS. Não obstante, os dispositivos devem emular essa demora.

O ambiente de experimentação apresentado permite que se verifique como há degradação de disponibilidade (Seção 8.7) quando se troca a topologia, ou seja, quando algum dispositivo entra no ou sai do modo dormente. O SustNMS efetua os cálculos necessários levando em consideração o lapso de tempo transcorrido para que o dispositivo se torne completamente operante. Tal tarefa é efetuada conforme o método REASoN (descrito e estendido no Capítulo 5).

Para o modelamento de tráfego, a solução adotada consiste no emprego do Controle de Tráfego do (LTC – *Linux Traffic Control*), que é uma funcionalidade baseada no núcleo do Linux. Essa funcionalidade permite emulação do gerenciamento da largura de banda dos enlaces. Por ter controle completo sobre os pacotes que atravessam o equipamento, o LTC possibilita o gerenciamento do tráfego de saída das NICs disponíveis. O método adotado para uso do LTC foi a disciplina de fila conhecida como *Hierarchical Token Bucket*, uma das disciplinas mais comuns para modelamento de tráfego (Gerdes e Bergmann, 2010).

Quanto à engenharia de tráfego, para que a solução oferecesse MPLS, uma abordagem baseada em *software* foi adotada, a saber, o projeto MPLS-Linux (*MPLS for Linux Project*, 2012). O sistema operacional utilizado é o GNU/Linux Debian, *kernel* 2.6.27.24. O ambiente resultante permite que a solução de gerenciamento energeticamente eficiente sendo experimentada possa forçar o roteamento que julgar mais adequado, sem que haja, entretanto, necessidade de implementação de um protocolo para tanto. Usar MPLS garante requisitos de QoS, como largura de banda, rotas explícitas, caminhos de reserva e rerroteamento rápido no caso de falhas. Usar MPLS no ambiente de experimentação tem o objetivo de facilitar a aplicação, por parte do sistema em teste, das rotas desejadas. A utilização de MPLS no ambiente de experimentação demandou interação com o sistema testado. Como exemplo de soluções verdes que forcem a rede a usar o caminho por elas selecionado, veja-se Seção 3.1.

Para permitir que o sistema testado inferisse informações quanto aos túneis, usou-se SNMP no ambiente de experimentação. O protótipo SustNMS beneficiou-se desse protocolo para acessar variáveis da MIB definida no RFC1213-MIB, a saber, *ifInOctets* e *ifOutOctets*. A partir do uso de comandos de linha (CLI – *Command-Line Interface*) transportados através da rede por SSH, o servidor do SustNMS define quais túneis os roteadores de ingresso devem configurar para os fluxos ingressantes. Pacotes SNMP e SSH constituem o sobrecusto gerado pelo sistema, pois são necessários na troca de mensagens do SustNMS, mas estariam ausentes numa operação sem este sistema.

7.3 Considerações do capítulo

Este capítulo contém a especificação dos requisitos de um ambiente de experimentação que permita a avaliação de sistemas de eficiência energética e apresentou uma possível implementação desse ambiente de experimentação. Tal implementação é baseada em roteadores virtuais implementados em *software* que operam sistemas com GNU/Linux. O sistema operacional dos roteadores deve ser alterado de modo que passem a fornecer serviços de gerenciamento, tais como suporte a SNMP, MIB, MPLS e controle de tráfego. Além dos requisitos próprios do ambiente de experimentação, a avaliação de uma solução verde requer outras entradas, como a topologia da rede onde a solução será avaliada, o perfil de tráfego e o perfil de consumo de energia elétrica dos equipamentos.

O Capítulo 8 apresenta como essas outras entradas foram utilizadas e discute os resultados obtidos quando o sistema de gerenciamento apresentado no Capítulo 6 é testado no ambiente de experimentação apresentado neste capítulo.

8 Descrição dos experimentos e seus resultados

Este capítulo apresenta os resultados experimentais obtidos a partir do protótipo do sistema descrito no Capítulo 6, quando este é testado no ambiente de testes descrito no Capítulo 7. O objetivo é avaliar os ganhos de energia obtidos por meio da atuação do SustNMS numa rede pequena, em oposição à degradação de qualidade de serviço (QoS) sentida na disponibilidade da rede, e ao sobrecusto impingido pela própria operação do sistema. A escalabilidade do sistema é abordada em termos de sobrecusto gerado pela solução na rede. Quatro experimentos com diferentes políticas de sustentabilidade são avaliados e comparados, mostrando como requisitos conflitantes e objetivos coalescem numa decisão que abarca o nível de QoS e ganhos de energia.

8.1 Políticas de sustentabilidade

Conforme mencionado acima, foram realizados quatro experimentos, cada qual com política distinta dos demais. O *primeiro experimento* (i) conta com uma política livre de restrições de eficiência energética ou confiabilidade e utiliza caminhos mínimos entre os clientes e servidores. O experimento (i) será também referido como o experimento-base, sendo o que mais precisamente reflete uma operação usual da rede, ou seja, uma operação desatenta ao consumo energético. O *segundo experimento* (ii) é composto por uma política que prioriza apenas requisitos de ganhos de energia e que, portanto, aceita certa degradação de QoS. No experimento (ii) espera-se que os ganhos de energia sejam máximos. A política do *terceiro experimento* (iii) assegura requisitos de desempenho da rede mais do que prioriza ganhos de energia. Tal política não aceita perda de pacotes até que a rede alcance os limites de sua capacidade. O *quarto experimento* (iv) tem uma política que assegura níveis de confiabilidade da rede mais do que

prioriza ganhos de energia. Todos os experimentos empregam a mesma topologia de rede, perfis de tráfego, perfis de consumo de energia e parâmetros de disponibilidade. Nessas políticas, os ganhos de energia são expressos por redução no consumo energético dos dispositivos. A confiabilidade da rede é medida segundo o método REASoN e o desempenho é medido em termos de pacotes perdidos.

8.2 Topologia da rede

A Figura 8 mostra a topologia da rede usada nos experimentos citados. A topologia é composta por cinco roteadores conectando quatro pontos de acesso. Dois destes pontos de acesso são servidores de vídeo, ao passo que os outros dois são clientes. Há caminhos redundantes quando todos os dispositivos estão ligados. Um plano de controle paralelo ao plano de dados conecta o SustNMS aos roteadores. A topologia empregada emula um cenário de distribuição de conteúdo.

Em cada experimento, o tráfego é gerado e roteado ao passo que os fluxos começam e sempre são priorizados caminhos de acordo com as restrições descritas na política correspondente. Podem-se enumerar quatro caminhos da topologia. Os caminhos empregados na alocação de tráfego originado em A3 são:

- caminho 1, correspondendo a (R1-R3-R4); e
- caminho 2, correspondendo a (R1-R2-R5-R4).

E os caminhos empregados na alocação de tráfego originado em A1 são:

- caminho 3, correspondendo a (R2-R3-R4); e
- caminho 4, correspondendo a (R2-R5-R4).

Figura 8 - Topologia integrada ao ambiente de experimentação

8.3 Perfil de tráfego

Os experimentos utilizam fluxos de vídeo sobre IP/MPLS como fonte de tráfego. O experimento é baseado em acesso intermitente de usuários consumidores de vídeos. O tráfego de entrada ultrapassa 30Mbps em dois momentos de pico. O perfil de tráfego usado consiste em quatro fluxos de vídeo que partem de A1 e A3. Os vídeos são consumidos por A2 e A4 (Figura 8). Os fluxos de vídeo são

iniciados assincronamente, como mostra a Figura 9. No máximo dois fluxos fluem simultaneamente do mesmo servidor de vídeo.

Figura 9 - Perfil de tráfego aplicado ao ambiente de experimentação proposto

8.4 Perfis de consumo

Perfis de consumo indicam como o consumo de energia de certo dispositivo varia em função da carga. Dois tipos de perfil de consumo foram utilizados:

- **Consumo linear com a carga:** Um tipo é baseado no comportamento descrito em (Antonakopoulos, 2010), de modo que a variação do consumo é linear com a carga, ou seja, quanto maior a carga, maior o consumo.

Naturalmente, tais equipamentos apresentam consumo mesmo quando a carga é nula, apenas por estarem ligados. Tal relação reflete um comportamento desejado e ambicionado em redes verdes (Bolla, 2011b). Assim, dispositivos com esse perfil representam equipamentos do futuro.

- **Consumo constante em função da carga:** O outro tipo de perfil de consumo utilizado não apresenta significantes variações com o aumento da carga, de sorte que o consumo seja mantido praticamente constante. Tal tipo de perfil representa perfis encontrados na literatura, como o que se vê em (Adelin, 2010). Equipamentos que contêm esse tipo de perfil representam equipamentos legados, comuns atualmente em redes de modo geral.

Como o ambiente de experimentação utiliza uma abordagem baseada em roteadores implementados em *software*, o sistema calcula o consumo de energia proporcionalmente à capacidade máxima do dispositivo físico correspondente. As capacidades máximas dos roteadores foram predefinidas como 30 Mbps para os roteadores R1, R2, R3 e R5, e como 40 Mbps para R4 (Figura 8). Como todo fluxo possível passa por R4 e nenhum outro roteador precisa lidar simultaneamente com todos os fluxos possíveis, os dispositivos foram parametrizados de modo que apenas R4 pudesse lidar com o tráfego máximo da rede. O tráfego máximo de 32 Mbps ocorre quando os quatro fluxos estão ativos. Por fim, os perfis de consumo devem incluir valores de consumo referentes ao modo dormente.

Quatro perfis de consumo são utilizados. Os perfis dos roteadores R1, R2, R4 e R5 variam linearmente e são baseados em (Bolla, 2011b). Eles representam equipamentos esperados no futuro e possuem diferentes fatores de escala, de sorte que emulem equipamentos variados. O roteador R3 tem consumo praticamente constante, representando equipamentos ineficientes e legados, conforme mostrado em (Adelin, 2010). As Equações 8.1 – 8.4 são usadas para indicar o consumo energético (em Watts) como função do uso. Nelas, a variável l representa a vazão sentida, e l_{max} representa a vazão máxima.

$$P_{R1} = P_{R4} = \begin{cases} 120, & \text{se dormente} \\ 200 + l \left(\frac{500}{l_{max}} \right), & \text{se ativo} \end{cases} \quad (\text{Equação 8.1})$$

$$P_{R2} = \begin{cases} 170, & \text{se dormente} \\ 200 + l \left(\frac{1000}{l_{max}} \right), & \text{se ativo} \end{cases} \quad (\text{Equação 8.2})$$

$$P_{R3} = \begin{cases} 250, & \text{se dormente} \\ 1000, & \text{se ativo} \end{cases} \quad (\text{Equação 8.3})$$

$$P_{R5} = \begin{cases} 220, & \text{se dormente} \\ 300 + l \left(\frac{5000}{l_{max}} \right), & \text{se ativo} \end{cases} \quad (\text{Equação 8.4})$$

8.5 Ganhos de energia

Os ganhos de energia obtidos pelo uso do SustNMS foram avaliados para as quatro políticas supracitadas. No experimento base, dispositivos que não estiverem lidando com carga também são mantidos ligados, ou seja, alguns dispositivos permanecem ociosos por algum tempo. A existência de dispositivos que às vezes estão ociosos permite que se pratique engenharia de tráfego e que se usem dispositivos redundantes como mecanismo de proteção contra falhas. A Figura 10 mostra a energia que seria consumida pelos diferentes caminhos *versus* a vazão, caso todo o tráfego da rede estivesse alocado num único caminho.

Figura 10 - Consumo de energia de cada caminho em diferentes condições de tráfego

Na Figura 10, o consumo energético por caminho mostra que a alocação ótima quanto a eficiência energética depende do tráfego. A figura mostra que, até 10Mbps de vazão, o consumo conjunto dos roteadores no caminho 1 é maior que o consumo conjunto dos roteadores no caminho 2. Além de 10 Mbps de vazão, o caminho 1 torna-se mais eficiente que o caminho 2. Deste modo, a questão que se segue é sobre quanta energia pode ser salva quando a alocação ótima para eficiência energética for priorizada.

Nos quatro experimentos, o SustNMS é utilizado na alocação do tráfego ilustrado na Figura 9. Fazendo-se isso, o sistema proposto aplica o Algoritmo 6.1 para determinar a alocação ótima de tráfego e quais os dispositivos que devem entrar em modo dormiente. A política descrita no experimento baliza a decisão do algoritmo. O comportamento da rede quando se aplica o Algoritmo 6.1 à topologia é visível nas Figura 11 e Figura 12.

No experimento (iv), a confiabilidade foi avaliada com base no método descrito por Amaral (2012), o qual mostra que, numa rede muito dinâmica, o impacto de se desligarem e ligarem frequentemente os equipamentos é expressivo, posto que se inclua nos cálculos o tempo que os dispositivos levam para acordar. Assim, toda vez que um dispositivo é posto no ou tirado do modo dormente, a confiabilidade da rede é reavaliada. A Seção 8.7 traz mais informações a este respeito.

Observa-se na Figura 11 que em todos os experimentos os roteadores R1, R2 e R4 estão sempre ligados, ou seja, tais roteadores nunca dormem. Isso ocorre porque fazer qualquer deles dormir implica desconexão na rede, o que não deve ocorrer. Por outro lado, os roteadores R3 e R5 são redundantes entre si e, dependendo do tráfego corrente, algum deles pode ser posto a dormir, mas jamais ambos. A figura mostra que, em cada experimento, tirar um roteador do modo dormente ou colocá-lo em tal modo resultam em comportamentos distintos. Tal fato expõe que, dependendo dos requisitos aplicados, mudanças na topologia podem ocorrer mais frequentemente.

A Figura 12 compara o consumo do experimento (i), não otimizado, com o consumo dos outros três experimentos. A figura mostra que ganhos de energia significativos são alcançáveis quando uma alocação mais eficiente é aplicada, como no caso do experimento (ii). A alocação aplicada no experimento (ii) conduz a 43% de ganhos de energia. Entretanto, os benefícios vêm com o custo de mensagens de controle adicionais e de uma inerente degradação de QoS da rede. O experimento (iii), o qual não tolera perda de pacotes, aloca o tráfego de modo a conduzir a 30% de ganhos de energia. O experimento (iv), o qual utiliza topologias lógicas de maior confiabilidade, alavanca os ganhos de energia a 27%. Os resultados mostram o compromisso existente entre apenas priorizarem-se ganhos de energia e considerarem-se também, em adição aos ganhos de energia, restrições de QoS.

Figura 11 - Mudanças aplicadas pelo SustNMS à topologia. Taxa de amostragem de 1 min.

Figura 12 - Consumo de energia de cada experimento

A taxa de amostragem também impacta o ganho total de energia. Se o sistema operar a altas taxas de amostragem, ele é capaz de mais precisamente perceber as mudanças que ocorrem no tráfego, no que tange à carga da rede. Uma taxa alta implica que o sistema pode tomar uma decisão sobre alocação de tráfego mais acurada do que no caso de uma taxa mais baixa. Tal comportamento está ilustrado na Figura 13.

Figura 13 - Consumo de energia agregado por taxa de amostragem, com ganhos de energia no destaque. Experimento (iii).

A Figura 13 mostra o consumo agregado do experimento (iii) *versus* diferentes taxas de amostragem. Na figura, as barras de base representam o consumo do experimento base. A figura expõe que, quanto mais frequente a amostragem é, tanto maiores são os ganhos de energia.

A seguir, encontram-se considerações sobre o sobrecusto gerado pelo sistema e sobre como o sobrecusto está correlacionado a ganhos de energia e à acurácia de se determinar a alocação ótima de tráfego. Na sequência, apresentam-se considerações sobre a confiabilidade do sistema estudado e uma avaliação do impacto sobre o tempo de vida dos equipamentos.

8.6 Sobrecusto de operação do sistema

O sobrecusto adicionado pela operação do sistema relaciona-se aos pacotes adicionais para medição e controle, os quais são necessários para coleta de dados e aplicação das decisões. Mensagens de medição são lançadas de acordo com a taxa de amostragem. A relação entre taxa de amostragem e quão frequentemente varia o tráfego e sua alocação determina a relevância dos dados coletados e a efetividade dos ganhos de energia, pois os caminhos ótimos são computados usando as informações de tráfego coletadas em tempo real. Para coletar os dados, o sistema requer um total de 2 pacotes SNMP (*Simple Network Management Protocol*) por amostra, dos quais um é de pedido e o outro, de resposta. Para forçar um roteador de ingresso a usar o túnel desejado, o sistema utiliza 1 pacote. Para definir em que estado de energia algum dispositivo deve estar, também o sistema utiliza 1 pacote. Este é um pacote de controle, contenedor de comandos CLI (*Command-Line Interface*). A presença dos supracitados tipos de pacotes foi capturada e confirmada pelo uso de filtros com o *sniffer* Tshark.

A Figura 14 compara o sobrecusto resultante quando se opera o SustNMS em cada experimento, a diferentes taxas de amostragem. O sobrecusto resultante relaciona-se às transições de topologia pelas quais a rede passa. Quando indo duma situação na qual apenas R5 está dormente para outra na qual apenas R3 está dormente, dois comandos CLI são despachados, dos quais um é para ativar R5, o outro para colocar R3 no modo dormente. A taxa de amostragem de 1 minuto, a despeito de impor mais sobrecusto, torna a rede mais verde, como mostra a Figura 15.

Figura 14 - Sobrecusto devido a pacotes de controle e monitoramento, durante 70 minutos, para os experimentos (ii), (iii) e (iv), alocado por taxa de amostragem

O sobrecusto de operação é menor quando uma taxa de amostragem menor é utilizada. Entretanto, taxas de amostragem menores têm o efeito colateral de diminuir a acurácia do sistema ao determinar a alocação do tráfego. A Figura 15 mostra a razão entre os ganhos de energia alcançada pelo SustNMS e o sobrecusto, para diferentes taxas de amostragem, para o experimento (ii). Da taxa de amostragem de 1 minuto até a taxa de amostragem de 10 minutos, a figura mostra uma tendência de diminuição dos ganhos de energia por sobrecusto. A acurácia das amostragens depende de como a última carga medida está relacionada à carga média do subsequente intervalo entre amostragens. Para taxas de 12 e 13 minutos, os instantes de amostragem não resultaram tão

representativos do correspondente intervalo entre amostragem. Assim, em tais casos o tráfego permaneceu subestimado por longos períodos, conduzindo a rede a perdas indesejadas e a altos ganhos de energia por sobrecusto. A Figura 15 mostra a energia salva, não a real qualidade dos ganhos, ou seja, se os ganhos vieram ao custo, por exemplo, de nenhum cliente ser atendido satisfatoriamente. Desta forma, apesar das taxas de 12 e 13 minutos apresentarem uma maior relação entre ganhos e sobrecusto, elas também apresentam maior perda de pacotes.

Figura 15 - Sobrecusto dividido pela redução de consumo, alocado por intervalo entre amostragens. Experimento (ii)

8.7 Confiabilidade

As ações impostas pelo SustNMS podem se realizar juntamente com um impacto inerente na confiabilidade da rede. Por este motivo, o sistema aplica uma avaliação dinâmica da confiabilidade quando ele precisa modificar o estado de energia de algum equipamento. Assim, o sistema pode garantir que os níveis mínimos desejados de confiabilidade sejam mantidos. Em Amaral (2012) mostrou-se que, em um cenário de rede muito dinâmico, é expressivo o impacto de frequentemente se colocarem ou tirarem os equipamentos do modo dormente, uma vez que nos cálculos estejam inclusos os lapsos temporais de se acordarem os equipamentos (mas confronte-se o Capítulo 5 desta dissertação, para uma extensão), estratégia que garante uma modelagem mais realista. No caso do método Reasoning, os impactos na confiabilidade são ainda mais expressivos, pois o cenário passa a ser semelhante a um cenário sem redundância. Como o modelo de Márkov usado para confiabilidade no método REASoN (Figura 3) é semelhante ao modelo usado para disponibilidade no método Reasoning (Figura 5), tem-se que a disponibilidade avaliada pelo método Reasoning, nos experimentos, corresponde à confiabilidade calculada pelo SustNMS (Figura 16).

Para avaliar tal impacto, cada vez que o estado da rede muda, o sistema calcula a confiabilidade condizente às próximas 24 horas, ou seja, a probabilidade da rede perfazer seus requisitos durante as 24 horas subsequentes à mudança de estado. Definiu-se o intervalo de 24 horas por ele representar a confiabilidade do intervalo de um dia e, logo, a de um ciclo do perfil de tráfego diário. Alternativamente, poder-se-ia utilizar um intervalo de análise de 8 horas, representando um turno de horas de trabalho. Outros parâmetros utilizados para o cálculo da confiabilidade são um MTTF de 60.000 horas, comumente encontrado em *data sheets* de fabricantes (CISCO, 2009), e um intervalo de *wake up* de 15 minutos. Esse intervalo representa o intervalo de tempo total para se ligar um dispositivo e aguardar até que a rede torne-se completamente operacional, após a ocorrência dalguma falha (CISCO, 2007).

Os resultados estão mostrados na Figura 16. Quando todos os equipamentos estão ligados, o SustNMS avalia a confiabilidade para 24 horas como $C_1(24h) =$

0,99999995. Quando apenas o roteador 3 está dormindo, o SustNMS avalia $C_2(24h) = 0,99998$. Quando apenas o roteador 5 está dormindo, o SustNMS avalia $C_3(24h) = 0,99999995$.

A diferença entre C_1 e C_3 não é significativa (ela ocorre apenas no décimo quinto dígito). Entretanto, C_2 difere das demais já no quinto dígito. As variações da confiabilidade da rede quando se opera o SustNMS, no experimento (iii), estão mostradas na Figura 16. Esse é o experimento que mais se assemelha às restrições de uma operação cotidiana, pois o experimento assegura desempenho antes de priorizar ganhos de energia. Com as extensões discutidas no Capítulo 5 espera-se que a diferença entre C_1 e C_3 seja substancialmente majorada.

Figura 16 - Avaliação da confiabilidade para cada mudança do estado da rede, nas situações de: 1) todos os dispositivos ligados; 2) apenas roteador 3 dormindo; e 3) apenas roteador 5 dormindo

8.8 Tempo de vida

Durante o experimento (iii) o roteador R3 entra uma vez em estado dormente e volta uma vez para o estado operacional (Figura 11). Já o roteador R5 repete esse comportamento 3 vezes. O experimento (iii) corresponde a 70 minutos de operação da rede. Supondo que durante um dia (24 horas) de operação esse comportamento seja repetido 16 vezes consecutivas, então o número total de ciclos diários do tipo modo dormente – modo operacional é 16 para R3 e 48 para R5. Suponha-se agora que nas demais horas (da madrugada) do período de 24 horas ambos os roteadores entrem em modo ocioso. Assim, a cada 24 horas cada roteador (R3 e R5) também perfaz 1 ciclo ocioso – ocupado. Sabe-se que, na ausência do sistema de gerenciamento energeticamente eficiente, R3 e R5 não entrariam no modo dormente e apenas teriam oportunidade de ficar ociosos naquelas mesmas horas da madrugada. Nesse cenário (esquematizado no Quadro 2), deseja-se saber o impacto da solução verde com respeito ao tempo de vida dos equipamentos.

Ciclos em 24 horas	Operação não sustentável		Operação sustentável	
	R3	R5	R3	R5
Tipo dormente – ativo	0	0	16	48
Tipo ocioso – ativo	1	1	1	1

Quadro 2 - Quantidade diária de cada ciclo térmico, por roteador, por tipo de operação

Seja *ciclo s* a denominação abreviada para um ciclo dormente – ativo e *ciclo o* a denominação abreviada para um ciclo ocioso – ativo. Seja ΔT_s a variação média de temperatura num ciclo s e ΔT_o a variação média de temperatura num ciclo o.

Pela Equação 4.4, o número médio de ciclos s até a falha é $MCTF_s = C_0(\Delta T_s)^{-q}$, e o número médio de ciclos o até a falha é $MCTF_o = C_0(\Delta T_o)^{-q}$, onde C_0 e q são constantes e, no caso em pauta, idênticas entre R3 e R5.

Se, no início do intervalo de 24 horas analisado, R3 estiver operando corretamente, então sua confiabilidade para um dia de operação sustentável é $R_{R3}(24h) = \exp(-(16 \text{ ciclos}/MCTF_s + 1 \text{ ciclo}/MCTF_o)) = \exp(-(16\lambda_s + \lambda_o) \times 24h)$, onde λ_s é a taxa de falhas ocasionada pelos ciclos s e λ_o é a taxa de falhas ocasionada pelos ciclos o , e a soma $(16\lambda_s + \lambda_o)$ é a taxa de falhas da operação sustentável. Ora, a taxa de falhas da operação não sustentável é λ_o . Logo, para R3 a relação entre as taxas de falhas dos dois tipos de operação é $AF_{R3} = (16\lambda_s + \lambda_o) / (\lambda_o) = 16.(\lambda_s/\lambda_o) + 1 = 16.[C_0(\Delta T_o)^{-q} / C_0(\Delta T_s)^{-q}] + 1 = 16.(\Delta T_s / \Delta T_o)^q + 1$.

Analogamente, para R5 a relação entre as taxas de falhas dos dois tipos de operação é $AF_{R5} = 48.(\Delta T_s / \Delta T_o)^q + 1$. Como $\Delta T_s, \Delta T_o > 0$, então $AF_{R3}, AF_{R5} > 0$ e a operação sustentável encurta o tempo de vida dos dois equipamentos. Na análise apresentada nesta seção foi considerado apenas o mecanismo de falhas em processadores ativado por ciclos térmicos, mecanismo o qual, dentre os apresentados, é o mais impactante. Outros mecanismos de falhas em processadores e mecanismos de falhas em ventoinhas são discutidos no Capítulo 4.

8.9 Considerações do capítulo

Este capítulo explorou o uso do ambiente de experimentação detalhado no Capítulo 7. Para tanto, empregou-se-o na avaliação do SustNMS, um sistema de gerenciamento de rede orientado por política de sustentabilidade. A utilização desse sistema pôs à prova a capacidade do ambiente de experimentação de fornecer e lidar com capacidades futuras pretendidas para equipamentos de rede. Mostraram-se resultados para quatro tipos de política, cada uma requerendo

capacidades específicas do ambiente e do sistema de gerenciamento. O ambiente de experimentação supre as necessidades do SustNMS, as quais representam também necessidades de sistemas de gerenciamento voltados a eficiência energética e, mais abrangentemente, a sustentabilidade.

Mostrou-se uma avaliação do impacto da solução verde no tempo de vida dos equipamentos. A avaliação focou nos efeitos dos ciclos térmicos e usou o perfil esperado de uso dos ciclos térmicos.

Os fatores que influenciam a eficiência alcançável são a taxa de amostragem, o perfil de tráfego, os perfis de consumo, o compromisso com o tempo de vida dos equipamentos e o compromisso com a disponibilidade ou confiabilidade da rede. A taxa de amostragem faz parte da configuração do sistema. O perfil de tráfego depende dos usuários da rede. Os perfis de consumo dependem da própria rede. Os compromissos com tempo de vida e disponibilidade ou confiabilidade dependem da política da rede, ulteriormente dependendo dos acordos de nível de serviço e dos administradores da rede.

9 Considerações finais

Nesta dissertação, foram propostos (i) um método para avaliação da confiabilidade e disponibilidade de redes cujos dispositivos utilizam outros estados energéticos como o modo dormente, (ii) um método para se avaliar o compromisso entre soluções de eficiência energética típicas e o tempo de vida dos equipamentos, método o qual também pode avaliar a viabilidade econômica da solução sustentável, e (iii) requisitos para obtenção de um ambiente de experimentação que viabilize a emulação de soluções de eficiência energética.

Os métodos foram aplicados para avaliação de um sistema de gerenciamento de rede orientado por política de sustentabilidade, mais especificamente eficiência energética, denominado SustNMS. A avaliação foi efetuada num ambiente de emulação de redes constituído por roteadores baseados em *software*. O sistema de gerenciamento é centralizado e reage a mudanças de carga da rede conduzindo o tráfego pelo caminho de comunicação mais eficiente, de acordo com políticas controladas pelo operador. Constatou-se que os ganhos de energia na presença de tráfego dinâmico são dependentes do quão frequentemente o estado da rede é observado pelo sistema de gerenciamento.

A avaliação experimental mostrou que ganhos de energia de até 43% são tangíveis quando se prioriza eficiência energética ao invés de uma solução de alocação de caminho mínimo e na qual os equipamentos ficam sempre ligados. Certa degradação no desempenho da rede e na confiabilidade é observada por uma avaliação em tempo real. Entretanto, experimentos seguintes mostraram que um compromisso controlável gera a proficuidade prática de ganhos de energia quando restrições de qualidade de serviço são definidas. Experimentos mostraram que 30% de economia de energia é alcançável quando nenhuma perda de pacote é permitida. Alternativamente, 27% de economia de energia é alcançado quando a confiabilidade calculada pelo SustNMS deve ser de “sete noves” (ou seja, entre 0,9999999 e 1). Tais ganhos são condicionados a uma taxa de prova de 1 minuto

e são adicionais aos ganhos inerentes a algoritmos que calculam caminhos energeticamente otimizados. Mediu-se o sobrecusto induzido pelo sistema de gerenciamento em termos de requisições sobre níveis de tráfego dos roteadores e de sinalização para o controle centralizado, a qual é necessária para se efetuarem trocas entre caminhos pré-configurados. Experimentos correlacionando a redução de energia alcançável às taxas de prova indicam como o custo de otimização se relaciona ao sobrecusto adicionado.

Construir um ambiente de experimentação para avaliar algoritmos de gerenciamento energeticamente eficientes é desafiador. Roteadores-padrão baseados em GNU/Linux carecem de várias funcionalidades importantes para emulação realista do comportamento da rede. Tais funcionalidades são a 1) configuração de perfis de consumo de energia, 2) emulação do estado dormente, 3) suporte a engenharia de tráfego, 4) suporte a máquinas virtuais com sistema operacional modificável, 5) controle de vazão, 6) mecanismo de aplicação de decisões e 7) provisão informação sobre perda de pacotes. Ainda, a abordagem desenvolvida, baseada em ambientes de experimentação, estende soluções existentes. No topo do *kernel* Linux, a solução integra emulação do estado dormente, suporte a MPLS e decisões de alocação baseadas em política. A abordagem foi validada por meio de experimentos que avaliaram o SustNMS, indicando a factibilidade do método para avaliar sistemas de gerenciamento de redes orientados a eficiência energética. Diferenças entre funcionalidades de *software routers* e *hardware routers*, tais quais o uso de perfis de consumo ao invés do uso de medições reais tomadas dos roteadores, são possíveis limitações do ambiente de teste, pois a acurácia da emulação depende dos parâmetros utilizados.

O método Reasoning, para cálculo de confiabilidade e disponibilidade, é inspirado no REASoN, mas utiliza uma modelagem mais acurada e algoritmos mais eficientes e temporalmente menos complexos. O modelo utilizado pelo Reasoning para cálculo de disponibilidade pode, nalgumas situações onde não há reparo, assemelhar-se formalmente ao modelo de confiabilidade do REASoN. Os modelos

de confiabilidade do Reasoning não têm pares no REASoN, pois temrinam já no estado de penalidade, pois este é considerado um estado de falha.

9.1 Contribuições

Esta dissertação contribui inicialmente com uma pesquisa sobre diferentes tipos de técnicas de eficiência energética em redes de computadores, descrevendo tanto técnicas locais a cada equipamento, quando técnicas que abrangem a rede toda, usualmente por meio de roteamento. É apresentada uma pesquisa sobre os mecanismos de falha física de equipamentos de rede, numa abordagem inicial para relacionar as técnicas verdes a uma degradação ou extensão do tempo de vida dos equipamentos.

Esta dissertação também explora os impactos da dormência de equipamentos, mecanismo utilizado por várias técnicas de eficiência energética, na confiabilidade e disponibilidade da rede. Tais impactos, por sua magnitude, merecem a atenção das operadoras de rede, as quais tipicamente estão mais interessadas na disponibilidade do serviço prestado do que na eficiência energética alcançável por técnicas verdes.

Por fim, discute-se aqui também como se obter um ambiente de experimentação para emulação de técnicas de eficiência energética. Muitas destas técnicas requerem funcionalidades ainda não completamente disponíveis nos equipamentos de rede, de modo que, para se emular uma solução verde, certos artifícios devem ser tomados.

9.1.1 Publicações

Esta dissertação está relacionada a outros quatro trabalhos dos quais seu autor tomou parte. São eles:

JANUARIO, G.C.; COSTA, C.H.A.; AMARAL, M.C.; RIEKSTIN, A.C.; CARVALHO, T.C.M.B.; MEIROSU, C., Evaluation of a policy-based network management system for energy-efficiency, 2013 IFIP/IEEE International Symposium on Integrated Network Management (IM 2013), Ghent, Belgium.

CARVALHO, T. C. M. B. ; MEIROSU, C. ; RIEKSTIN, A. C. ; AMARAL, M. ; JANUARIO, G. C. ; COSTA, C. H. A. ; MIERS, C. ; GABOS, D. ; CHENG, E. L. ; FIGUEIREDO, L. B. Sustainability-Oriented Policies applied to Network Management - A practical view of refinement and application of sustainability-oriented policies. In: 2013. (Org.). Mini-Course in SBRC 2013. 1ed.SBRC: SBRC, 2013, v. 1, p. 259-318.

CARVALHO, T. C. M. B.; RIEKSTIN, A. C.; AMARAL, M. C.; COSTA, C. H. A.; JANUARIO, G. C.; DOMINICINI, C. K.; MEIROSU, C. Towards Sustainable Networks Energy Efficiency Policy from Business to Device Instance Levels. In: Proceedings of the 14th ICEIS. 2012, Wroclaw, Poland.

COSTA, C. H. A ; AMARAL, M. C. ; JANUÁRIO, GUILHERME C. ; CARVALHO, T. C. M. B. ; MEIROSU, C. SustNMS: Towards Service Oriented Policy-Based Network Management for Energy-Efficient Networks. In: Proceedings of the 2nd IEEE/IFIP Conference on Sustainable Internet and ICT for Sustainability, 2012.

Em (Januário, 2013) descreve-se como se obter um ambiente de experimentação para avaliação de soluções de eficiência energética. Em (Carvalho, 2013) estudam-se soluções e técnicas de eficiência energética e métodos para refinamento de políticas que rejam tais soluções e técnicas, sendo o refinamento, ele mesmo, uma tal solução. Em (Carvalho, 2012) estuda-se uma arquitetura para refinamento em cinco camadas de políticas, com aplicação a sustentabilidade em redes. Em (Costa, 2012) descreve-se uma arquitetura para gerenciamento de redes orientado por política e voltado para eficiência energética. A arquitetura permite que as políticas avaliem o compromisso com qualidade de serviço e disponibilidade da rede.

9.2 Trabalhos Futuros

Neste capítulo identificam-se pontos desta dissertação que podem ser estendidos. Os pontos comentados correspondem a alguns, mas provavelmente não a todos,

estudos com potencial para divulgação científica. Os pontos de extensão, organizados por áreas do conhecimento em soluções para redes mais sustentáveis, são:

Tempo de vida – Um maior desenvolvimento do viés econômico da avaliação sobre os impactos da operação sustentável no tempo de vida dos equipamentos possibilitaria a criação de funções de utilidade que abarquem diferentes tipos de compromisso. Nota-se ainda que as taxas de falhas dos mecanismos estudados no Capítulo 4 e, de modo mais notável, na Seção 4.2.4, podem ser, a princípio, aproveitadas em um modelo de Márkov que as avalie em conjunto com os efeitos da demora em acordar do modo dormiente.

Confiabilidade e disponibilidade – Pode-se estudar qual o passo que levaria os Algoritmos 5.4 e 5.3 (para cálculo iterativo da confiabilidade e disponibilidade quando a demora em despertar do modo dormiente é contabilizada), à convergência e desenvolver modelos que juntem tanto a dormência de enlaces e dispositivos, quanto também suas taxas de falha. Buscar na literatura, ou mesmo obter por medição própria, dados estatísticos relativos à demora em despertar dos dispositivos possibilitaria precisar quão boa é a aproximação da ocorrência dessa demora a uma distribuição exponencial.

Ambiente de experimentação - Pode-se-o estender pela adição de um arcabouço de política que possibilite o refinamento automático de política, ou que viabilize políticas dinamicamente variáveis. O refinamento de política ideal deveria partir do nível de negócios, passando pelos níveis de rede, de dispositivo e de instância. Dinamicidade refere-se tanto à mudança dos valores dos parâmetros da política, quanto à própria inclusão de novos parâmetros. Tal arcabouço contribuiria para a redução da complexidade do gerenciamento de redes, tornando este mais simples do ponto de vista de gerenciamento.

REFERÊNCIAS

Advanced Configuration and Power Interface Specification; Hewlett-Packard Corporation, Intel Corporation, Microsoft Corporation, Phoenix Technologies Ltd., Toshiba Corporation; Revision 4.0a, April 5, 2010. Sítio: <http://www.acpi.info/DOWNLOADS/ACPIspec40a.pdf> (acessado em 15 de Janeiro de 2014)

Adelin, A.; Owezarski, P.; Gayraud, T. “On the impact of monitoring router energy consumption for greening the internet,” in Grid Computing (GRID), 2010 11th IEEE/ACM International Conference on, oct. 2010, pp. 298 –304.

Altıparmak, F.; Dengiz, B.; Smith, A. “Reliability estimation of computer communication networks: Ann models,” in **Computers e Communication, 2003. (ISCC 2003). Anais. Eighth IEEE Inter- national Symposium on**, june-3 july 2003, pp. 1353 – 1358 vol.2.

Amaral, M. C. et al. “REASoN - REliability and/or availability evaluation for sustainable networking,” in 4th Int. Work. on Reliable Networks Design and Modeling (RNDM’12), St. Petersburg, Russia, Oct. 2012.

Antonakopoulos, S.; Fortune, S.; Zhang, L. “Power-aware routing with rate-adaptive network elements,” in GLOBECOM Workshops (GC Wkshps), 2010 IEEE, dec. 2010, pp. 1428 –1432.

Beller, A.; Jamhour, E.; Pellenz, M. Defining Reusable Business-Level QoS Policies for DiffServ

Black, J. R. Electromigration 8212;a brief survey and some recent results. Electron Devices, IEEE Transactions on, 16(4):338 – 347, apr 1969.

Blanquicet, F.; Christensen, K. An Initial Performance Evaluation of Rapid PHY Selection (RPS) for Energy Efficient Ethernet.2007

Blanquicet, F. "PAUSE Power Cycle: A new backwards compatible method to reduce energy use of Ethernet switches," White paper, Ethernet Alliance, Apr. 2008.

Bolla, R. et al. "Energy efficiency in the future internet: A survey of existing approaches and trends in Energy-Aware fixed network infrastructures," IEEE Communications Surveys & Tutorials, vol. 13, no. 2, pp. 223–244, 2011.

Bolla, R. et al. "The potential impact of green technologies in next-generation wireline networks: Is there room for energy saving optimization?" Communications Magazine, IEEE, vol. 49, no. 8, pp. 80 –86, august 2011.

Boros. Policy-based network management with SNMP, **Anais do EUNICE**, p. 13–15. University of Twente, Holanda, set. 2000.

Bussab, W. O. Estatística básica. Saraiva, 2008.

Callou, G. et al. "Sustainability and dependability evaluation on data center architectures," Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on , vol., no., pp.398,403, 9-12 Oct. 2011.

Carvalho, T. C. M. B. C. Modelo para gestão de TIC orientada à política de sustentabilidade. Tese de Livre Docência, Escola Politécnica – Universidade de São Paulo, São Paulo, 2012.

Chabarek, J; Barford, P. "Power-awareness extensions for network testbeds," in Communications Workshops (ICC), 2011 IEEE International Conference on, june 2011, pp. 1 –6.

Chandramouli, M. et al. "Power and Energy Monitoring MIB, Network Working Group, IETF, Tech. Rep., Jan 2012.

Chaudhari, S. et al. "Green-IT: An approach to energy savings using energy aware network management system," in Communications (NCC), 2012 National Conference on, feb. 2012, pp. 1 –5.

Cianfrani, A. et al. "An Energy Saving Routing Algorithm for a Green OSPF Protocol," INFOCOM IEEE Conference on Computer Communications Workshops , 2010 , vol., no., pp.1,5, 15-19 March 2010

CISCO "Cisco resilient ethernet protocol," Cisco, USA, Tech. Rep., 2007.

Costa, C. H. A. et al. "SustNMS: towards service oriented Policy-Based network management for Energy-Efficiency," in Second IFIP Conference on Sustainable Internet and ICT for Sustainability (SustainIT 2012) (SustainIT 2012), Pisa, Itália, Out. 2012a.

Costa, C. H. A. et al. "Dynamic Method to Evaluate Code Optimization Effectiveness" in **Anais do 15th International Workshop on Software and Compilers for Embedded Systems**, St. Goar, Germany, 2012b, pp. 62–71.

Dasgupta, A.; Karri, R. Electromigration reliability enhancement via bus activity distribution. In **Design Automation Conference Anais** 1996, 33rd, pages 353 – 356, jun, 1996.

Davy, S.; Jennings, B.; Strassner, J. "The policy continuum - a formal model," **Anais do Second IEEE International Workshop on Modelling Autonomic Communications Environments**, MACE, 2007, pp. 65–79.

Dongmei, W.; Guangzhi, L. "Efficient distributed bandwidth management for mpls fast reroute," Networking, IEEE/ACM Transactions on, vol. 16, no. 2, pp. 486 –495, april 2008.

Emulab. The Emulab Testbed. <http://www.emulab.net>, 2010.

Gerdes, S.; Bergmann, O. "Building a test environment for emulating link characteristics of disruptive networks," in Wireless Communications and Networking Conference (WCNC), 2010 IEEE, april 2010, pp. 1 –5.

GeSI "Smart 2020: Enabling the low carbon economy in the information age," The Climate Group on behalf of the Global eSustainability Initiative (GeSI), Tech. Rep., Jun. 2008.

GeSI. SMARTer 2020: The Role of ICT in Driving a Sustainable Future. Technical Report. GESI. Brussels, Belgium, 2012.

Ghein, L. "MPLS Fundamentals". Cisco Press. 21 Nov. 2006

Green, H.; Hant, J.; Lanzinger, D. "Calculating network availability," in 2009 IEEE Aerospace conference. IEEE, Mar. 2009, pp. 1–11.

Gunaratne, C.; Christensen, K.; Nordman, B. "Managing Energy Consumption Costs in Desktop PCs and LAN Switches with Proxying, Split TCP Connections, and Scaling of Link Speed," Int'l J. Network Management, vol. 15, no. 5, pp. 297-310, Sept./Oct. 2005.

Gunaratne, C. et al. Reducing the Energy Consumption of Ethernet with Adaptive Link Rate (ALR).

Gupta, M.; Grover, S.; Singh, S.. A Feasibility Study for Power Management in LAN Switches

He, F.; Qi, H., "A method of estimating network reliability using an artificial neural network," in Pacific-Asia Workshop on Computational Intelligence and Industrial Application, 2008. PACIIA '08, vol. 2. IEEE, Dec. 2008, pp. 57–60.

Heller, B. et al. ElasticTree: saving energy in data center networks. **Anais do 7th USENIX conference on Networked systems design and implementation(NSDI'10)**. USENIX Association, Berkeley, CA, USA, 17-17.

IETF. Policy Framework Architecture, 1999. Disponível em <http://tools.ietf.org/html/draft-ietf-policy-arch-00>.

Januário, G. C. et al. Evaluation of a Policy-Based Network Management System for Energy-Efficiency. In [submitted for publication], September 2012.

JEDEC. Failure mechanisms and models for semiconductor devices. Technical Report JEP122C, Solid State Technology Association - JEDEC, 2006.

Jin, X. et al. Physics-of-failure analysis of cooling fans. In Prognostics and System Health Management Conference (PHM-Shenzhen), 2011, pages 1 –8, may 2011.

Johnson, B. W. The Design and Analysis of Fault Tolerant Digital Systems. Addison-Wesley, Jan. 1989.

Juniper. “Limited Lifetime Warranty EX2200, EX3200, EX4200”, 2011. Disponível em: www.juniper.net/support/warranty/990235.pdf. Último acesso: Mar., 2013

Juniper. EX3200 Ethernet Switches, 2013. Disponível em: www.juniper.net/us/en/local/pdf/datasheets/1000216-en.pdf. Último acesso: Mar., 2013

Kempf, J. et al. “OpenFlow MPLS e the open source label switched router,” in **Anais do 23rd International Teletraffic Congress**, ser. ITC '11. ITCP, 2011, pp. 8–14.

Klutke, G. A.; Kiessler, P.C.; Wortman, M.A.. A critical look at the bathtub curve. Reliability, IEEE Transactions on, 52(1):125 – 129, march 2003.

Lam, Y. F.; Li, V. O. K. “A survey of network reliability modeling and calculations,” in IEEE Military Communications Conference - Communications-Computers: Teamed for the 90's, 1986. MILCOM 1986, vol. 1. IEEE, Oct. 1986, pp. 1.2.1–1.2.5.

Lange, C. et al. “Energy consumption of telecommunication networks and related improvement options,” IEEE Journal of Selected Topics in Quantum Electronics, vol. 17, no. 2, pp. 285–295, Apr. 2011.

Lee, U. et al. Toward energy-efficient content dissemination. IEEE Network, 25(2):14–19, April 2011.

Lin, C. et al. “A mesh network reliability analysis using reliability block diagram,” in 2010 8th IEEE International Conference on Industrial Informatics (INDIN). IEEE, Jul. 2010, pp. 975–979.

Maciá-Pérez, F.; Marcos-Jorquera, D.; Gilart-Iglesias, V. “Energy management system as an embedded service: Saving energy consumption of ICT,” in **Anais do 22nd International Conference on Architecture of Computing Systems**, ser. ARCS '09. Berlin, Heidelberg: Springer-Verlag, 2009, pp. 195–206.

Manral, V. “Benchmarking power usage of networking devices,” Network Working Group, IETF, Tech. Rep., May 2010.

Miers, C. Uma Arquitetura Usando Trackers Hierárquicos para Localidade em Redes P2P Gerenciadas. Tese de Doutorado, Escola Politécnica – Universidade de São Paulo, São Paulo, 2012.

Mostowfi; Christensen. Saving Energy in LAN Switches: New Methods of Packet Coalescing for Energy Efficient Ethernet Mehrgan Mostowfi e Ken Christense

Mpls for Linux Project, 2012. Disponível em: <http://mplslinux.sourceforge.net/>

Neto; Cymbalista. Probabilidades. Editora: Edgard Blucher. 2006.

Oltsi; Jon. A Business-Driven Approach to Mobile Enterprise Security; White Paper. Editora: Juniper, 2012.

Pecht, M.; Radojcic, R.; Rao, G. Guidebook for Managing Silicon Chip Reliability. CRC Press, 1 edition, December 1998.

Rana, A.I.; O Foghlu, M.; , "New Role of Policy-Based Management in Home Area Networks - Concepts, Constraints and Challenges," New Technologies, Mobility and Security (NTMS), 2009 3rd International Conference on , vol., no., pp.1-6, 20-23 Dec. 2009.

Rana, A.I. et al. "Autonomic Policy-based HAN Traffic Classification using Augmented Meta Model for Policy Translation". In **Anais do IEEE and IFIP International Conference on wireless and Optical communications Networks (WOCN 2011)** , May 2011, Paris, France.

Rosing, T.S.; Mihic, K.; De Micheli, G., "Power and Reliability Management of SoCs," Very Large Scale Integration (VLSI) Systems, IEEE Transactions on , vol.15, no.4, pp.391,403, April 2007

Rubio-Loyola, J. et al. A Functional Solution for Goal-oriented Policy Refinement, 2006.

Schrijver, A. (1998). Theory of linear and integer programming. John Wiley and Sons, 1998. ISBN 978-0-471-98232-6.

Shooman, M. L. Reliability of Computer Systems and Networks: Fault Tolerance, Analysis, and Design, 1st ed. Wiley-Interscience, Dec. 2001.

Sidharth; Sundaram, S. A methodology to assess microprocessor fan reliability. In Thermal and Thermomechanical Phenomena in Electronic Systems, 2004. ITherm '04. The Ninth Intersociety Conference on, pages 96 – 101 Vol.1, June 2004.

Skiena; Revilla. Programming Challenges: The Programming Contest Training Manual, Springer-Verlag, New York, 2003.

Sorensen, S. The Sustainable Network: The Accidental Answer for a Troubled Planet. O'Reilly Media, 1 edition, October 2009.

Srinivasan, J. et al. Ramp: A model for reliability aware microprocessor design. Technical Report RC23048, IBM, White Plains, NY, IBM Research, 2003.

Srinivasan, J.; et al. The case for lifetime reliability-aware microprocessors. In Computer Architecture, 2004. **Anais. 31st Annual International Symposium on**, pages 276 – 287, June 2004.

Strassner, J. "How policy empowers business-driven device management," Policies for Distributed Systems and Networks, 2002. Proceedings. Third International Workshop on, vol., no., pp.214,217, 2002

Strassner, J. Policy Based Network Management: solutions for the next generation. Elsevier, EUA, 2003

Takeshita, H. et al. "Demonstration of the self organized dynamic link power management by MiDORi; energy optimal network topology design engine," in OFC/NFOEC, March 2011, pp. 1 –3.

The Network Simulator. <http://www.isi.edu/nsnam/ns>, 2010.

Tian, X.. Cooling fan reliability: failure criteria, accelerated life testing, modeling and qualification. In Reliability and Maintainability Symposium, 2006. RAMS '06. Annual, pages 380 –384, jan. 2006.

Trivedi, K. S.; Selvamuthu, D., “Markov modeling in reliability,” in Encyclopedia of Quantitative Risk Analysis and Assessment. John Wiley & Sons, Ltd, 2008.

Wilson, D. S.; Wilkinson, J. W. Reliability model for miniature blowers per MIL-B-23071B, Final Report Feb. 1974 - Mar. 1975. Technical report, Shaker Research Corporation, 1975.

Wu, E et al. Interplay of voltage and temperature acceleration of oxide breakdown for ultra-thin gate oxides. Solid-state Electronics Journal, 2002.

Yeh, W. “A particle swarm optimization approach based on monte carlo simulation for solving the complex network reliability problem,” IEEE Transactions on Reliability, vol. 59, no. 1, pp. 212–221, Mar. 2010.

Zhang, M. et al. “GreenTE: Power-aware traffic engineering,” in **Anais do 18th IEEE International Conference on Network Protocols**, ser. ICNP '10. Washington, DC, USA: IEEE Computer Society, 2010, pp. 21–30.

Zill; Cullen. Equações Diferenciais, 3a Ed., Makron Books, 2001.