Este exemplar corresponde à redação final da tese defendida por GIOVANNI MOURA DE HOLANDA e aprovada pela Comissão Julgadora em 20 / 05 / 92

Orientador

Avaliação da Segurança de Funcionamento como Garantia da Qualidade de Serviço: Aplicações em Sistemas de Telecomunicações

TESE DE MESTRADO SUBMETIDA À
FACULDADE DE ENGENHARIA ELÉTRICA
DA UNIVERSIDADE ESTADUAL DE CAMPINAS

Autor: Giovanni Moura de Holanda

20 de Maio de 1992

Orientador: Édson Moschim
Co-orientador: Jorge Moreira de Souza

Banca Examinadora

Membros Titulares:

- Prof. Dr. Édson Moschim, Faculdade de Engenharia Elétrica da UNICAMP.
- Prof. Dr. Ivanil Bonati, Faculdade de Engenharia Elétrica da UNICAMP.
- Prof. Dr. Edmundo A. Souza e Silva, Núcleo de Computação Eletrônica da UFRJ.

Membro Suplente:

- Dr. Jorge Moreira de Souza, CPqD-TELEBRÁS.

A Tê e a Paula

Agradecimentos

Gostaria de expressar meus agradecimentos ao CPqD-TELEBRÁS, cujo apoio institucional viabilizou a realização deste trabalho. Agradeço às equipes dos projetos SAMSAT e COMPAC, que colaboraram, da melhor maneira possível, no levantamento das informações necessárias à avaliação destes dois sistemas.

Sou igualmente grato a Ivanil Bonati e Edmundo Souza e Silva por terem aceitado integrar a banca examinadora, bem como por suas valiosas e oportunas sugestões.

Meus agradecimentos a José Francisco M.S. Franco, por sua participação na análise do SAMSAT, e a Marta R.B. Martini, pela cotidiana e enriquecedora troca de idéias. Agradeço a Édson Moschim pela confiança e pelo apoio, e, muito particularmente, agradeço a Jorge Moreira de Souza que, além de ter sido o principal incentivador deste empreendimento, percorreu comigo cada etapa do mesmo.

Aproveito também para agradecer a José e Criselda Holanda, meus pais, pelo apoio incondicional que sempre deram às minhas iniciativas.

Enfim, gostaria de transmitir a minha gratidão àqueles que contribuíram, com incentivos, sugestões e subsídios de qualquer natureza, para a consolidação deste trabalho.

Índice

Resumo

Nos últimos anos, é possível notar um interesse crescente por sistemas de telecomunicações capazes de fornecer serviços com um alto nível de qualidade e de segurança de funcionamento. Isto decorre do fato de que os usuários de telecomunicações dependem cada vez mais de sistemas com capacidade de operar de forma contínua e livre de falhas.

Os sistemas de telecomunicações devem, portanto, ser desenvolvidos com o propósito de atender aos requisitos de segurança de funcionamento impostos por este cenário emergente. Para tanto, o seu desenvolvimento deve ser acompanhado de um processo de avaliação que permita verificar o atendimento a estes requisitos e, quando preciso, reorientar o desenvolvimento rumo às metas estabelecidas.

Neste contexto, o objetivo desta dissertação compreende a abordagem prática de técnicas e ferramentas utilizadas atualmente na avaliação dos atributos de segurança de funcionamento, com enfoque voltado para as falhas de origem hardware. O processo de avaliação, juntamente com a aplicação de algumas destas técnicas, é ilustrado através de dois sistemas desenvolvidos pelo CPqD-TELEBRÁS: o SAMSAT, um sistema AMDT de comunicação de dados via satélite; e o nó de comutação de pacotes (PSN) do sistema COMPAC.

Além disso, a avaliação da segurança de funcionamento destes dois sistemas inclui dois aspectos importantes: *i*) uma proposta de metodologia para levantamento do fator de cobertura, a qual foi empregada no SAMSAT; e *ii*) no caso do PSN COMPAC, uma análise de como um procedimento operacional de comutação periódica, realizado entre unidades ativa e reserva deste nó, pode aumentar o nível de segurança de funcionamento do sistema.

Introdução Geral

A vertiginosa evolução da eletrônica tem sido responsável pelo marcante desenvolvimento dos sistemas de telecomunicações nas últimas decadas. O setor de telecomunicações vem crescendo incessantemente, tanto no que se refere ao atendimento a uma demanda em franca expansão quanto à multiplicidade de serviços e aplicações, de tal forma que é praticamente impossível detectar indícios de desaceleração para o futuro próximo. Pelo contrário, é possível preconizar novas aplicações em áreas cada vez mais variadas da sociedade atual.

Em todos os ramos da atividade, a qualidade do serviço prestado aos usuários dos sistemas de telecomunicações tem conquistado níveis crescentes de importância, principalmente em função da necessidade de padrões técnicos mais rígidos, advindos das inovações tecnológicas, e do aumento do nível de exigência dos usuários em termos de qualidade. Neste aspecto, uma filosofia de qualidade de serviço tem sido esboçada no sentido de acompanhar a evolução destes sistemas e garantir condições para o fornecimento de serviços com alto nível de qualidade.

Recomendando no âmbito das telecomunicações, o CCITT [CCI 84], [CCI 88] define a Qualidade de Serviço como o efeito coletivo do desempenho de serviço, o qual determina o grau de satisfação do usuário e o quanto o funcionamento do sistema se aproxima do ideal. A percepção desta qualidade de serviço pelo usuário é determinada por fatores de desempenho que se relacionam entre si, tais como: facilidades de operação, desempenho de tráfego, segurança de funcionamento, desempenho de transmissão, etc.

A relação entre estes fatores pode ser observada quando da ocorrência de uma falha no sistema. Este evento pode conduzir a uma situação onde o sistema tem seu funcionamento prejudicado, como consequência de uma incapacidade total de operação ou de uma degradação parcial do desempenho de tráfego e/ou de transmissão, ocasionando transtornos como: custos de manutenção, congestionamento, insatisfação do usuário, perdas de receita, etc. Neste aspecto, a segurança de funcionamento assume um papel de suma importância, dado que é constituída por um conjunto de atributos que caracterizam o comportamento do sistema mediante a ocorrência de falhas (por exemplo, capacidade do sistema funcionar sem falhas durante um determinado tempo, capacidade do sistema se encontrar operacional quando o serviço é solicitado, etc.), cujas medidas são representadas pelos parâmetros confiabilidade, disponibilidade, tempo médio entre falhas, etc.

Em função de tal relevância, os requisitos associados aos parâmetros de segurança de funcionamento devem ser cuidadosamente especificados durante o desenvolvimento do sistema, de forma a proporcionar o nível de qualidade de serviço adequado a cada aplicação. O atendimento a estes requisitos é propiciado através da utilização de técnicas de prevenção e de tolerância a defeitos, e da avaliação dos parâmetros de segurança de funcionamento . Na fase de desenvolvimento do sistema, a avaliação fornece insumos necessários à validação dos requisitos de segurança de funcionamento, fomentando, quando preciso, a tomada de

decisões que possam reorientar o projeto no sentido de contemplar as especificações.

Usualmente, o processo de avaliação da segurança de funcionamento de um sistema envolve quatro fases: *i*) escolha do parâmetro que se comporta como o melhor indicativo da segurança de funcionamento do sistema, *ii*) análise das características operacionais e funcionais do sistema, *iii*) modelagem, e, por fim, *iv*) obtenção e análise das estimativas do parâmetro escolhido. Para que a realização destas fases proporcione resultados satisfatórios, torna-se fundamental o emprego adequado de técnicas e ferramentas de avaliação.

Neste contexto, o objetivo desta dissertação compreende a abordagem prática de técnicas e ferramentas utilizadas na avaliação da segurança de funcionamento, com aplicações em dois estudos de caso de sistemas de telecomunicações. Estas técnicas e ferramentas visam, sobretudo, proporcionar avaliações capazes de prever o comportamento do sistema em termos dos atributos de segurança de funcionamento e, assim, fornecer subsídios valiosos ao desenvolvimento de sistemas em consonância com os requisitos de qualidade de serviço propostos. A avaliação é baseada em modelos de descrição de estados, com enfoque direcionado para as falhas de origem hardware. No caso de sistemas com arquitetura e princípios de funcionamento complexos, estes modelos descrevem o comportamento operacional do sistema segundo um processo markoviano.

Desta maneira, o conteúdo do trabalho aqui apresentado está estruturado do seguinte modo: O capítulo 1, introdutório, focaliza a avaliação da segurança de funcionamento como expediente para a obtenção do nível de qualidade de serviço requerido pelas aplicações às quais os sistemas de telecomunicações se destinam, justificando, em termos gerais, a escolha das técnicas e ferramentas utilizadas na avaliação. Adicionalmente, são apresentados os conceitos e terminologia de segurança de funcionamento empregados nos capítulos posteriores, através de uma correspondência entre o significado dos termos e o contexto no qual se enquadram ao longo deste trabalho.

O capítulo 2 discorre sobre as principais técnicas e ferramentas utilizadas atualmente na modelagem dos sistemas de telecomunicações, trilhando uma sequência que cobre desde noções básicas do processo markoviano, passando por modelos de Markov para avaliação dos parâmetros de segurança de funcionamento (com ênfase para a confiabilidade e disponibilidade) e por métodos práticos de resolução para modelos complexos, até uma breve descrição sobre o panorama geral de ferramentas e técnicas de última geração.

O capítulo 3 traz uma aplicação prática destas técnicas de avaliação no estudo de caso do sistema SAMSAT, um sistema AMDT de comunicação de dados/voz via satélite, focalizando as quatro fases do processo de avaliação, e apresentando a metodologia empregada no levantamento do fator de cobertura de falhas de suas estruturas redundantes. Esta metodologia foi concebida com a finalidade de prover modelos de segurança de funcionamento mais realistas, com capacidade de representação mais próxima das condições reais de funcionamento do sistema.

O capítulo 4 mostra uma outra aplicação, desta vez referente à avaliação da segurança de funcionamento de um nó de comutação de pacotes do sistema COMPAC. Através desta

aplicação é possível constatar um exemplo típico de como os resultados obtidos com a avaliação podem apontar alternativas que contribuam para aumentar o nível de segurança de funcionamento de um sistema, e, deste modo, viabilizar o desenvolvimento de sistemas em consonância com os requisitos de qualidade de serviço.

Finalmente, são apresentadas as conclusões de caráter geral concernentes ao assunto abordado, incluindo perspectivas futuras de expansão e aprofundamento deste trabalho.

Capítulo 1

Segurança de Funcionamento: Conceitos Básicos

Os equipamentos eletrônicos que compõem os sistemas de telecomunicações estão sujeitos a falhas geradas por eventos de diversas naturezas, tais como, componentes com problemas de fabricação, projetos inadequados, condições ambientais anormais, etc. A ocorrência destes problemas afeta, de maneira direta, alguns fatores de desempenho determinantes da qualidade do serviço prestado pelo sistema, conforme mencionado na introdução geral. Isto contribui para a insatisfação do usuário, e proporciona transtornos de ordem operacional e financeira para a empresa operadora do serviço. Tais decorrências tornam-se críticas em se tratando de sistemas que atuam sob demanda e com compartilhamento de recursos entre um grande número de usuários, o que requer o funcionamento correto (dentro das especificações) e, se possível, ininterrupto de seus equipamentos.

Além disso, a evolução dos sistemas no cenário de telecomunicações tem levado à adoção de padrões técnicos mais rígidos, principalmente se considerada a utilização de técnicas digitais, e tem sido acompanhada por um aumento gradativo do nível de exigência dos usuários [GRU 86]. Como resposta a estas tendências, inclusive às exigências operacionais de determinadas aplicações, eleva-se a necessidade de fornecimento de serviços com alto grau de segurança de funcionamento, de modo que o sistema atenda aos requisitos de qualidade impostos por este cenário emergente [RIC 88].

A garantia de desenvolvimento de sistemas com o nível de segurança de funcionamento almejado é viabilizada através de procedimentos como, por exemplo, a avaliação dos atributos de segurança de funcionamento. Com o propósito de situar este tipo de avaliação no contexto esboçado pelos procedimentos para garantia de qualidade de serviço, e de introduzir , de uma forma geral, os métodos utilizados para obtenção de medidas dos principais atributos de segurança de funcionamento, este capítulo é constituído pelos seguintes tópicos: A seção 1.1 apresenta as características conceituais dos atributos de segurança de funcionamento; A seção 1.2 mostra a terminologia e conceitos associados à segurança de funcionamento, os quais são empregados ao longo deste trabalho; A seção 1.3 destaca os principais métodos de tolerância a defeitos, com o intuito de apresentar os mecanismos e procedimentos que são frequentemente mencionados nos capítulos subsequentes; A seção 1.4 discorre sobre os métodos gerais de avaliação da segurança de funcionamento,

apontando os fatores que condicionam a escolha das técnicas utilizadas neste trabalho; A seção 1.5 apresenta as medidas associadas aos atributos de segurança de funcionamento, relacionando-as de acordo com os objetivos de segurança de funcionamento inerentes ao tipo de aplicação do sistema.

1.1 Segurança de Funcionamento

A segurança de funcionamento (*dependability*) representa o nível de confiança que pode, justificadamente, ser depositado no serviço fornecido pelo sistema [LAP 90]. De acordo com o CCITT [CCI 88], a segurança de funcionamento de um sistema de telecomunicações (o que abrange redes, equipamentos, etc.) tem seus objetivos representados pelo desempenho conjunto de atributos como disponibilidade, confiabilidade, mantenabilidade e suporte de manutenção, como pode ser observado na taxonomia apresentada pela figura 1.1.

A confiabilidade, no seu sentido mais amplo, corresponde à capacidade do sistema desempenhar uma função requerida dentro de um intervalo de tempo definido . No sentido métrico, ou seja, como medida de segurança de funcionamento, esta capacidade é representada em termos de probabilidade. Já a disponibilidade traduz a capacidade do sistema desempenhar uma função requerida num dado instante de tempo ou em qualquer instante dentro de um dado intervalo de tempo. Similarmente, a disponibilidade como medida de segurança de funcionamento é definida por uma representação probabilística desta capacidade. As definições matemáticas destas duas medidas são apresentadas no capítulo 2.

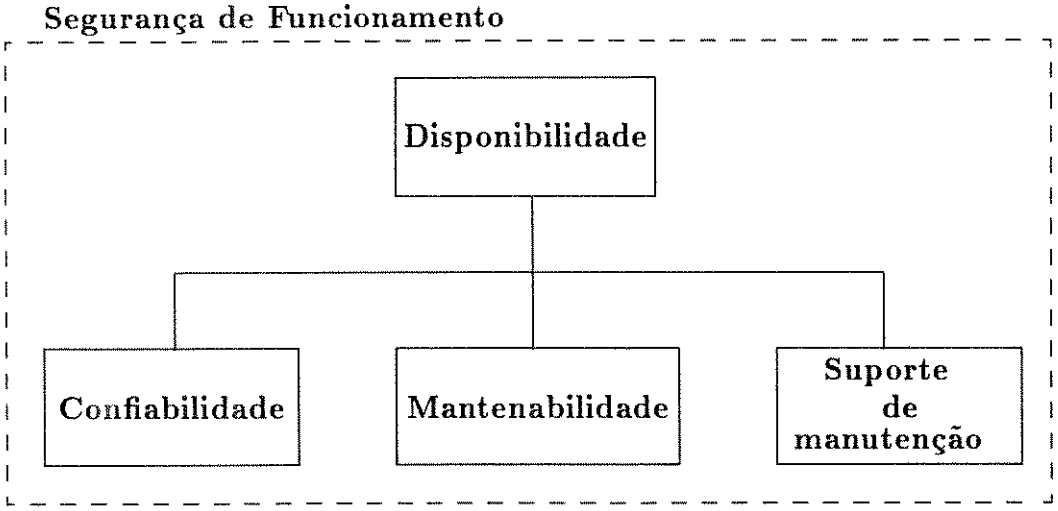


Figura 1.1 - Taxonomia de segurança de funcionamento.

A mantenabilidade representa a capacidade do sistema retornar a um estado no qual possa desempenhar corretamente suas funções, quando a manutenção é desempenhada sob dadas condições e usando procedimentos e recursos estabelecidos. O suporte de manutenção é a capacidade de uma organização responsável pela manutenção fornecer, em

função da demanda e segundo uma dada estratégia de manutenção, os recursos necessários à manutenção do sistema.

É importante observar que a disponibilidade dos sistemas de telecomunicações está calcada no efeito coletivo do próprio desempenho de confiabilidade, mantenabilidade e suporte de manutenção, e que apesar do condicionamento existente entre confiabilidade e disponibilidade de um sistema é possível a existência de diferenças significativas entre estas duas medidas. Este último aspecto pode ser justificado através de particularidades relativas à natureza intrínseca de cada medida, ou seja, um sistema com alta incidência de falhas, consequentemente baixa confiabilidade, pode ter alta disponibilidade, desde que possa ser restaurado rapidamente, conforme mencionado em [LEI 87].

A distinção entre estas medidas surge como consequência direta dos objetivos e particularidades inerentes ao serviço fornecido pelo sistema, os quais ditam a necessidade de abordagens específicas que permitam focalizar os atributos de segurança de funcionamento mais adequados a cada aplicação. Sistemas não assistidos (e.g., equipamentos situados em locais de difícil acesso, carga útil de satélite, sistemas embarcados de telemetria, etc.) requerem confiança e continuidade de funcionamento, portanto é razoável qualificar a confiabilidade como o principal atributo de segurança de funcionamento neste campo de aplicação. Por outro lado, em se tratando de sistemas reparáveis cujo não funcionamento prejudica um grande número de usuários ou acarreta em congestionamentos e perdas de receitas (e.g., controle e telessupervisão de sistemas elétricos, centrais de serviços telemáticos, equipamentos pertencentes a redes de telecomunicações, etc.), mas que assim mesmo admitem pequenos intervalos sem fornecimento de serviços, a disponibilidade se impõe como o principal requisito de segurança de funcionamento.

Nestes casos, os equipamentos devem possuir uma alta segurança de funcionamento, de forma que o sistema atenda aos requisitos operacionais e econômicos condizentes com o tipo de aplicação a que se destina. Para garantir tais objetivos, alguns procedimentos como prevenção de defeitos, tolerância a defeitos, estratégia de reparos e avaliação de segurança de funcionamento são utilizados no desenvolvimento/fabricação de equipamentos eletrônicos.

O primeiro conjunto de procedimentos, classificados como prevenção de defeitos, envolve o controle da qualidade de produção, técnicas de *screening*, revisões de projeto, etc. A tolerância a defeitos envolve a utilização de arquiteturas redundantes e mecanismos de supervisão de falhas, os quais tornam possível o processo de reconfiguração entre as unidades ativa e reserva, mediante a ocorrência de falhas.

A estratégia de reparos reflete as facilidades de suporte de manutenção disponíveis, e possibilita o retorno do sistema/equipamento a uma condição operacional. Além disso, é importante ressaltar que a duração do reparo tem influência marcante na confiabilidade/disponibilidade dos sistemas redundantes e reparáveis.

A necessidade de aferir a eficiência do emprego de todos estes procedimentos, juntamente com os aspectos econômicos e as decisões associadas aos compromissos de pro-

jeto, requer a quantificação dos atributos da segurança de funcionamento e a utilização de métodos e ferramentas que possibilitem avaliá-los.

Antes de abordar propriamente as técnicas de avaliação da segurança de funcionamento e de ilustrá-las através de estudos de caso, conforme o objetivo deste trabalho, faz-se necessário a apresentação de alguns conceitos e terminologia adotada no decorrer dos capítulos subsequentes, bem como uma rápida descrição dos principais métodos de tolerância a defeitos usualmente empregados.

1.2 Conceitos e Terminologia Associada

À medida que uma determinada área de conhecimento evolui, estabelecendo conteúdo próprio e bem definido, há, naturalmente, a formação de um jargão técnico associado ao seu campo de investigação. No tocante à segurança de funcionamento dos equipamentos eletrônicos, este jargão tem tomado formas a partir de um conjunto de conceitos e definições que, inclusive, refletem o processo evolutivo e o tipo de aplicação dos próprios equipamentos.

Primeiramente, durante a conquista de um mercado consumidor, era desejável que os equipamentos funcionassem durante o tempo requerido pelo usuário, de forma que a confiabilidade correspondia à principal propriedade de segurança de funcionamento a ser focalizada. Com a subsequente obtenção de equipamentos/sistemas confiáveis, os serviços se multiplicaram e o mercado passou por um quadro de expansão e diversificação, criando uma demanda por serviços fornecidos prontamente, de acordo com a solicitação dos usuários. Neste particular, a disponibilidade se torna o melhor indicativo do atendimento a estas necessidades.

Adicionalmente, a utilização de equipamentos em aplicações críticas [LAP 86] e a disseminação crescente dos sistemas computacionais distribuídos têm fomentado a adoção dos conceitos de segurança contra riscos (*safety*) e contra intrusões (*security*) como atributos da concepção geral de segurança de funcionamento [LAP 90]. O surgimento do conceito de segurança contra riscos advem do fato dos sistemas terem atingido um nível de confiança que os tornam qualificados a aplicações onde uma falha pode trazer consequências catastróficas. Por sua vez, a sedimentação do conceito de segurança contra violações é justificada pelo fato de que, mesmo com a obtenção de sistemas computacionais seguros do ponto de vista de riscos, algumas aplicações apresentam-se vulneráveis a intrusões, ou seja, defeitos introduzidos "conscientemente" no sistema com o propósito de conduzi-lo a uma situação de falha, cujas consequências podem ser catastróficas. Não obstante, aplicações com tais características não chegam a integrar significativamente o cenário dos sistemas de telecomunicações, de maneira que estes dois conceitos fogem ao escopo deste trabalho.

Como pode ser observado, a abordagem das diferentes nuances de segurança de funcionamento condiciona a definição de diferentes atributos (que devem ser expressos através de medidas de segurança de funcionamento) , os quais, contudo, são complementares e

interrelacionados. Além disso, a intenção de atingir um elevado nível de segurança de funcionamento leva à utilização de técnicas e métodos capazes de assegurar este intento, o que, por sua vez, gera uma terminologia associada. É importante notar que o alto grau de subjetividade inerente a alguns conceitos de segurança de funcionamento implica na formulação de termos e definições nem sempre consensuais, tornando conveniente o explicitamento da terminologia adotada.

Assim, este capítulo apresenta a terminologia empregada subsequentemente, através de um formato que procura fornecer um mapeamento entre o significado dos termos e seu respectivo contexto. Imbuído deste propósito, o conteúdo apresentado a seguir diz respeito aos conceitos e propriedades dos entraves à segurança de funcionamento, i.e., falhas, erros e defeitos. Por motivos didáticos, os conceitos e terminologia associados aos métodos de tolerância a defeitos e às técnicas gerais de avaliação da segurança de funcionamento são apresentados nos itens subsequentes (1.3 e 1.4, respectivamente). Em qualquer caso, todos os conceitos e terminologia correspondem aos principais componentes do campo de investigação delineado pelo tema proposto neste trabalho.

1.2.1 Entraves à segurança de funcionamento

Os entraves à segurança de funcionamento de um sistema, segundo a taxonomia apresentada em [LAP 90], são caracterizados pela existência de falhas, erros e defeitos. Em outras palavras, estão associados a circunstâncias não desejadas que podem causar ou serem resultadas de eventos responsáveis por uma degradação na segurança de funcionamento de um determinado sistema.

Os termos **falha, erro** e **defeito**, quando utilizados no contexto de segurança de funcionamento de sistemas, possuem significados diferentes, apesar de apresentarem uma certa interdependência fenomenológica. Conceitualmente, uma **falha** representa a impossibilidade de um sistema[1] desempenhar as suas funções mediante o surgimento de **erros** (no sistema ou no ambiente em que está inserido), os quais são ocasionados pelos mais variados tipos de **defeitos**.

Os processos de origem e manifestação das falhas, erros e defeitos exemplificam este processo de interdependência. Segundo [LAP 90], um defeito é considerado **ativo** quando ele produz um erro, caso contrário, encontra-se num estado **dormente** com possibilidade de tornar-se, ou não, ativo. Um erro pode ser **latente**, enquanto não for reconhecido como tal, ou não for **detectado** por um mecanismo de detecção. Um erro pode desaparecer antes de ser detectado, ou pode propagar-se, originando outros erros. Durante a operação do sistema, a presença de um defeito ativo só pode ser determinada pela detecção de um erro. Já uma falha ocorre quando um erro afeta o serviço fornecido pelo sistema e é percebido pelo usuário. A falha de um componente resulta em um defeito para o sistema que contém o componente.

[1] Este termo é aqui empregado no sentido de **caixa preta**, uma entidade com comportamento externo bem definido e interagindo com outras entidades dentro de um determinado ambiente. Neste sentido, um usuário pode ser considerado como uma entidade situada no referido ambiente.

Um exemplo elucidativo deste processo pode ser obtido a partir de um simples defeito hardware no sistema. A ocorrência de um curto-circuito num Circuito Integrado (CI) constitui uma falha, em termos de especificação do circuito. A nível de funcionamento do sistema ao qual o CI pertence, as consequências proporcionadas pelo curto-circuito (e.g., bit grampeado num valor fixo) caracterizam um defeito que permanecerá dormente enquanto não for ativado. Uma vez ativo (a partir da solicitação do componente onde reside o defeito), o defeito produz um erro (e.g., uma sequência de bits errada ou um estado lógico com valor diferente do que deveria ser) que pode afetar o serviço prestado pelo sistema, caracterizando uma falha.

No que concerne aos defeitos, suas fontes e mecanismos de manifestação são bastante diversificadas. Em função disso, e do fato de sua ocorrência representar o "epicentro" de um desencadeamento de eventos que podem levar a uma quebra no fornecimento do serviço, a seguir é apresentada uma compilação das principais propriedades dos defeitos. Esta compilação é baseada nas abordagens descritas em[NEL 90] e [LAP 90], onde, neste último, os defeitos são classificados segundo sua natureza, persistência e origem.

Do ponto de vista de sua natureza, é possível dividir os defeitos em dois grupos: **defeitos acidentais**, os quais aparecem ou são criadas ao acaso, e os **defeitos intencionais**, os quais são criados deliberadamente.

A origem dos defeitos pode ser estratificada segundo três aspectos: causas, limites de sistema, e fase de criação dentro do ciclo de vida do sistema. Por sua vez, estes aspectos apresentam os seguintes desdobramentos:

- Em termos das causas, os defeitos podem ser divididos em físicos e provocados pelo homem. Os **defeitos físicos** são resultantes de fenômenos físicos adversos, enquanto os **defeitos provocados pelo homem** resultam de imperfeições inerentes à condição humana
- Em termos dos limites de sistema, os defeitos podem ser considerados como **internos**, correspondendo às partes do sistema que podem produzir um erro quando ativadas, e **externos**, os quais resultam de interferências ou interações com o ambiente físico, por exemplo, pertubações eletromagnéticas, vibração, temperatura, etc.
- Em termos da fase de criação, os defeitos podem ser diferenciados em: **defeitos de projeto**, resultantes de imperfeições ocorridas durante o desenvolvimento do sistema ou de modificações subsequentes; e **defeitos operacionais**, os quais ocorrem durante a operação comercial do sistema.

Quanto à persistência temporal, os defeitos podem ser categorizados em permanentes, transientes e intermitentes [NEL 90]. Os **defeitos permanentes** são condições resultantes de falhas de componentes ou erros de projeto, as quais não são corrigidas com o tempo. Os **defeitos transientes** são normalmente provocados por disturbios externos esporádicos e ocorrem durante uma quantidade limitada de tempo. Os **defeitos intermitentes** são originados a partir de condições de operação instável dos componentes (e.g., mudanças

nos parâmetros de um componente, devidas a capacitâncias parasitas que surgem e desaparecem reiteradamente). Desta forma, um sistema com defeito intermitente tem seu funcionamento oscilando entre condições livres de defeitos e condições com a presença de defeitos.

Com este panorama das propriedades dos defeitos, é possível identificar os principais modos de falha, aos quais os sistemas estão sujeitos. Os modos de falha têm sido exaustivamente analisados ao longo de muitos anos, a partir de testes do ciclo de vida de componentes e de dados de falha coletados em campo. Uma quantidade significativa de dados de falhas tem sido compilada em manuais publicados regularmente, como por exemplo, a série MIL-HDBK-217, cujas versões mais recentes são [MIL 86] e [MIL 91], respectivamente.

As observações realizadas durante o ciclo de vida de um componente/sistema possibilitam conclusões importantes em termos de caracterização dos modos de falha com relação ao tempo de vida. No começo de vida dos sistemas, ou de suas partes, há a ocorrência de uma grande quantidade de **falhas prematuras**, devidas, na maior parte, a defeitos físicos decorrentes do processo de fabricação. Vencido este período, o sistema entra numa fase caracterizada por uma incidência menor de falhas. Entretanto, as causas destas falhas são difíceis de determinar, uma vez que normalmente ocorrem em função de eventos aleatórios como variações ambientais e stress excessivo. Como consequência, as falhas associadas a este período são denominadas de **falhas aleatórias**. Após este período, e à medida que o sistema envelhece, suas partes passam por um processo de deterioração e, novamente, muitas falhas ocorrem. As falhas transcorridas neste período são comumente denominadas de **falhas por desgaste**.

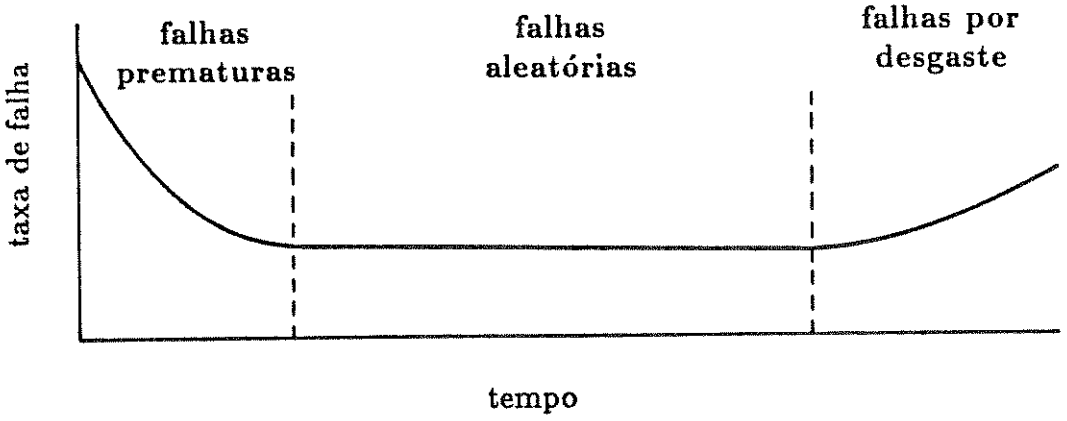


Figura 1.2 - Curva da banheira.

Uma ilustração típica destes três modos de falhas é proporcionada pela curva da banheira, apresentada na figura 1.2. Esta curva mostra a relação entre a taxa de falha (λ) de um sistema e o seu tempo de vida, onde é importante observar que a parte plana

da curva, correspondendo a região de falhas aleatórias, permite, no caso de componentes eletrônicos, aproximar a taxa de falha deste período para um valor constante. Um aspecto importante decorrente desta condição é que o período caracterizado por uma taxa de falha constante predomina em termos de vida útil do sistema, tornando possível a hipótese de sistemas/componentes com taxa de falha constante. Esta hipótese simplifica enormemente o processo de modelagem, como pode ser melhor percebido na abordagem apresentada no capítulo 2.

1.3 Métodos de Tolerância a Defeitos

Os **métodos de tolerância a defeitos** compreende um conjunto de mecanismos e procedimentos cujo objetivo é possibilitar o funcionamento do sistema diante das eventuais presenças de defeitos. Em se tratando dos equipamentos eletrônicos, a obtenção de um sistema com arquitetura hardware tolerante a defeitos está calcada no emprego de redundância e de mecanismos de tolerância a defeitos que proporcionem a utilização eficiente das partes redundantes. Neste aspecto, a especificação e implementação destes mecanismos constituem meios para a obtenção de sistemas com o nível de segurança de funcionamento desejado.

O impacto da utilização de princípios de tolerância a defeitos na segurança de funcionamento de um sistema pode ser percebido, em termos qualitativos, através de uma relação entre a probabilidade de ocorrência de defeitos e de falhas. Utilizando a relação apresentada em [NEL 90] e aplicando-a à probabilidade de funcionamento de um sistema (P_{func}), é possível escrever:

$$P_{func} = P\{\text{não ocorrência de defeitos}\} + P\{\text{operação correta}|\text{ocorrência de um defeito}\}.P\{\text{ocorrência de um defeito}\}$$

Nesta equação, o primeiro termo representa a probabilidade de não ocorrência de defeitos, a qual depende dos aspectos de prevenção de defeitos, tais como: componentes de alta qualidade e metodologias formais de projeto. O segundo termo representa a probabilidade de ocorrência de um defeito que, por sua vez, não conduz a uma falha do sistema. Em outras palavras, este termo representa o aumento da probabilidade de funcionamento com o emprego de métodos e procedimentos de tolerância a defeitos.

É importante notar, ainda com relação a esta equação, que $P\{\text{operação correta} \mid \text{ocorrência de um defeito}\}$ representa a probabilidade condicional de que o sistema continue em operação dado a ocorrência de um defeito. Portanto, reflete a capacidade de cobertura associada aos mecanismos de tolerância a defeitos.

A **cobertura de defeitos** está associada aos mecanismos de supervisão, e reflete a capacidade de um sistema redundante recuperar-se prontamente quando da ocorrência de falhas em suas unidades[2]. A recuperação do sistema reflete a correta utilização da re-

[2]É importante notar que, do ponto de vista de funcionamento do sistema, a falha de uma unidade constitui um defeito deste sistema. Tal defeito, quando não coberto, leva o sistema a uma condição de falha.

dundância, a qual está atrelada à eficiência dos mecanismos de supervisão de defeitos, i.e., unidades de deteção/localização de defeitos e unidades de comutação/reconfiguração das partes redundantes. Em termos quantitativos, a cobertura representa a probabilidade destes mecanismos funcionarem corretamente, dado que uma falha ocorreu, e tem influência direta na confiabilidade/disponibilidade do sistema. Esta influência, bem como os aspectos de modelagem da cobertura, tem sido largamente documentada [ARN 73], [GEI 83], [TRI 84], [MCG 85], [BAZ 86], [DUG 89].

Todos estes mecanismos fazem parte da estratégia de tolerância a defeitos, e compreendem:

Detecção de erro: É constituído por mecanismos hardware e software, e corresponde à identificação de um estado errôneo, a partir dos sintomas proporcionados por um defeito.

Detecção de defeito: Está associado à detecção de um erro. Na verdade, o processo de detecção atua sobre os sintomas do defeito, o que torna o termo detecção de erro mais preciso e abrangente. Entretanto, é comum na literatura e nas especificações técnicas de sistema, a utilização do termo detecção de defeitos. Muitos mecanismos são normalmente empregados na detecção de erros/defeitos, tais como: paridade de bits, teste de consistência, violação de protocolos, circuitos "vigilante"(watch dog timer) etc. Segundo o modo de atuação no sistema, estes mecanismos podem ser enquadrados em duas categorias: detecção on-line e detecção off-line. A **detecção on-line** engloba facilidades para detecção em tempo real, ou seja, enquanto o sistema desempenha suas funções. Por outro lado, a **detecção off-line** não capacita o sistema a funcionar durante a realização dos testes de detecção, atuando antes da operação ou em intervalos especificados.

Mascaramento: Correção dinâmica dos erros gerados, a partir de técnicas que escondem os seus efeitos. Um exemplo típico destas técnicas é a utilização de votação por maioria em arquiteturas redundantes. É possível também a existência de um mascaramento de defeitos na ausência de erros (e.g., uma alternância periódica entre as unidades de um sistema redundante pode "mascarar" um defeito, ainda não ativado, numa destas unidades).

Localização de defeitos: é constituído por mecanismos implementados em harware e software, os quais possibilitam a identificação da parte defeituosa, responsável pelo erro detectado. Dependendo da abrangência dos mecanismos de localização de defeitos, a identificação da parte defeituosa pode ocorrer a nível de módulo[3], placa[4] ou componente[5], conforme a estrutura hierárquica/funcional do sistema.

[3]Estrutura física que comporta um conjunto de funções ou subfunções reunidas de forma a refletir a metodologia de projeto. Em termos de hierarquização, a composição ordenada de vários módulos forma um equipamento.

[4]Corresponde a um conjunto de componentes agrupados com o intuito de realizar funções específicas. Um módulo pode ser formado por uma ou mais placas.

[5]Corresponde ao elemento de mais baixo nível hierárquico. Numa avaliação sistemica do hardware, representa, por exemplo, os componentes eletro-eletronicos, peças mecanicas, etc.

Reconfiguração: Após a detecção de um defeito, o sistema pode proceder a uma substituição do item[6] (componente, placa, módulo, etc.) defeituoso. O item pode ser substituído por equivalentes reservas, através de mecanismos automáticos ou manuais. Adicionalmente, o item pode ser apenas bloqueado e o sistema operar de forma degradada, constituindo o processo de **degradação suave.**

Recuperação: Quando os efeitos de um erro são eliminados, e.g., após uma reconfiguração, o sistema retoma a sua operação normal. Para sistemas computacionais, alguns mecanismos são implementados com o propósito de recuperar o processamento e permitir o funcionamento do sistema a partir do ponto em que o defeito foi detectado.

Reiniciação: Determinadas situações impossibilitam o processo de recuperação, por exemplo, quando um erro provoca a perda de uma grande quantidade de informação. Além disso, o sistema pode não possuir facilidades de recuperação. Em ambos os casos, o sistema pode ser projetado com mecanismos de reiniciação, os quais possibilitam o reinício das operações. Em função dos danos provocados pelo defeito, a reiniciação pode ser total (em todas as operações, inclusive a partir do ponto em que o defeito foi detectado) ou parcial (preservando apenas alguns processos). Não obstante, algumas situações impossibilitam a retomada das operações a partir do ponto em que a falha foi detectada, conduzindo à perda das tarefas em andamento.

Diagnose: Os mecanismos referentes ao processo de detecção/localização de defeitos podem fornecer, além do desempenho normal de suas funções, sinalização externa com o objetivo de orientar a atividade de manutenção. Estes mecanismos de sinalização representam um aspecto importante em termos de mantenabilidade do sistema, porém não cobrem todas as situações. Neste caso, procedimentos adicionais de diagnose são requeridos no sentido de prover informações sobre a localização do defeito, o que, inclusive, viabiliza a aplicação de uma estratégia de reparo eficiente.

Reparo: Compreende a substituição ou manutenção de um item defeituoso. O processo de reparo pode ocorrer com o sistema ativado ou desativado. Com o sistema ativado, ou seja, em funcionamento normal, o item defeituoso pode ser substituído imediatamente por um reserva, enquanto é reparado, obedecendo a um procedimento similar ao da reconfiguração, ou o sistema pode continuar seu funcionamento porém sem o item defeituoso. Outra possibilidade de reparação com o sistema ativado ocorre nos casos de mascaramento por redundância e de degradação suave, onde o sistema pode continuar operando sem o item defeituoso. A reparação com o sistema desativado ocorre nos casos em que todo o sistema deve sair de operação para propiciar os procedimentos de diagnose e reparo.

Mecanismos de supervisão: Correspondem ao conjunto formado pelos mecanismos de detecção, localização e reconfiguração.

No que diz respeito à utilização de estruturas hardware redundantes, os sistemas usualmente têm sua arquitetura implementada segundo dois tipos de configuração: estática e

[6] O termo item pode ser usado para designar qualquer elemento da estrutura hierárquica

dinâmica [SIE 84].

Uma configuração **estática** ou **passiva** é utilizada no mascaramento de um determinado número de defeitos, sendo adequada para sistemas não assistidos e que requerem um elevado nível de segurança de funcionamento por um curto intervalo de tempo.

Exemplos deste tipo de configuração são as redundâncias série, paralelo, bimodal (série-paralelo), votador por maioria (esta configuração é caracterizada por uma estrutura redundante paralela, onde decisões são tomadas por um votador que compara sinais provenientes dos itens em paralelo com sinais remanescentes. As decisões são feitas apenas quando o número de itens úteis é maior do que o número de itens em falha, caso contrário, a estrutura é considerada em falha.), etc.

Segundo [SIE 84], a **redundância estática** tolera defeitos mas não os sinaliza, ou seja, emprega mecanismos de mascaramento de defeitos sem, entretanto, incorporar mecanismos de detecção. Neste sentido, o mascaramento é uma forma de redundância estática, onde as interconexões lógicas providas pelo hardware permanecem fixas, sem a ocorrência de intervenções externas. Desta maneira, quando esta redundância é exaurida pela presença de defeitos, qualquer defeito adicional provocará erros e consequentemente a falha do sistema.

A **redundância dinâmica** envolve a reconfiguração dos itens do sistema diante o surgimento de falhas, ou o roteamento alternativo quando se trata de enlaces de comunicação. Este tipo de redundância é indicada para aplicações que exigem um elevado nível de segurança de funcionamento por um longo período de tempo, e admitem procedimentos de reparo.

De uma forma geral, tanto o mascaramento como a detecção de defeitos são usados como parte da redundância dinâmica. (Neste aspecto, [NEL 90] distingue as configurações redundantes em três tipos: estática, dinâmica e híbrida. Na híbrida, é tomada como base uma configuração estática que pode mascarar um dado número de defeitos, enquanto os itens defeituosos são detectados e reparados.)

Uma técnica de redundância dinâmica muito empregada é a redundância **standby** [KUM 80], [OSA 76], onde itens redundantes de uma estrutura em paralelo podem ser comutados entre circuitos ativo e reserva, uma vez que os itens possuem mecanismos de detecção e têm suas saídas conectadas a um polo comutador. Para sistemas com processamento em tempo real, uma forma usual de redundância standby é a **hot standby**, a qual consiste na operação simultânea de todos os itens redundantes. Neste tipo de configuração, como pode ser observado na figura 1.3, um item ($A_{n=1,\ldots,N}$) se encontra na condição ativa enquanto o(s) outros(s) se encontra(m) na reserva, podendo passar à condição ativa mediante detecção de falha do primeiro. Neste caso, as funções são transferidas automaticamente para um item reserva que se econtra sincronizado com os demais, sem interrupção de funcionamento do sistema.

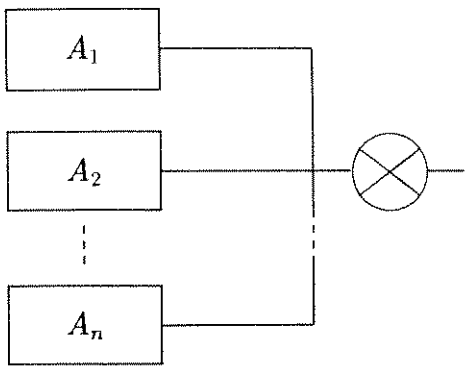


Figura 1.3 - Redundância *hot standby*.

É importante notar que os mecanismos de detecção formam a base da redundância dinâmica [SIE 84], uma vez que as chances de uma reconfiguração ocorrer com sucesso depende fortemente da capacidade de detecção de defeitos. Esta probabilidade de sucesso é comumente usada no processo de avaliação das medidas de segurança de funcionamento.

1.4 Métodos de Avaliação

No que diz respeito aos métodos de avaliação dos parâmetros de segurança de funcionamento, é possível destacar duas grandes categorias: medição e modelagem. Na categoria de medição, os parâmetros podem ser obtidos através da observação do comportamento do sistema em campo (operação comercial ou condições reais de funcionamento) [STR 88] ou através de simulação baseada em protótipos ou emuladores e de métodos de Monte Carlo [SHO 68]. Entretanto, no que tange à avaliação preliminar de alternativas de projeto, ou mesmo à avaliação de sistemas que ainda não acumularam experiência de funcionamento em campo, estas técnicas apresentam-se como ferramentas pouco adequadas. Neste particular, a categoria de modelagem constitui-se numa opção viável, permitindo avaliações que refletem as possibilidades de representação do sistema, e conciliando aspectos como estágio de desenvolvimento e recursos disponíveis com precisão e tempo de obtenção dos resultados pretendidos.

Este tipo de avaliação é extremamente útil, em especial para os projetistas e gerentes de projeto, e pode ser viabilizado através de modelos de confiabilidade[7] capazes de:

- Prever a confiabilidade/disponibilidade de sistemas

[7]Neste caso, o termo confiabilidade é empregado no ***sentido lato***, correspondendo a uma designação genérica que se confunde com o próprio domínio conceitual de segurança de funcionamento. Desta forma, modelo de confiabilidade compreende a denominação dada aos modelos com capacidade de fornecer as medidas de segurança de funcionamento, onde confiabilidade e disponibilidade são as mais usuais, e os modelos de confiabilidade viabilizam esta determinação. A designação confiabilidade foi aqui empregada com o intuito de alertar sobre a possibilidade deste termo ser usualmente encontrado na literatura com esta conotação, mesmo em detrimento ao uso crescente do conceito segurança de funcionamento. Entretanto, daqui em diante o termo adotado em tal contexto será segurança de funcionamento.

- Identificar pontos que levam a confiabilidade/disponibilidade a níveis desejados, permitindo o reprojeto de partes ou subsistemas ainda na fase de desenvolvimento do produto.
- Verificar o atendimento aos requisitos de confiabilidade/disponibilidade.
- Fornecer subsídios para a escolha de estratégias de reparo que atendam às necessidades dos usuários e fornecedores de serviço.
- Constituir alternativas aos testes de ciclo de vida do produto, os quais demandam grandes quantidades de recursos técnicos (e consequentemente econômicos) e tempo de teste elevado.

No processo de modelagem, três tipos de modelos são normalmente utilizados: modelos de contagem por partes, modelos combinatórios e modelos de descrição de estados [REI 91]. Os modelos de contagem por partes proporcionam análises onde o sistema é decomposto em unidades funcionais ou subsistemas, de tal forma que a falha de uma destas unidades caracteriza a falha do sistema.

Já os modelos combinatórios, considerados como extensões dos primeiros, incluem outras possibilidades de descrição do comportamento operacional do sistema. Usualmente o sistema é decomposto em subsistemas ou unidades que, para efeito de análise, podem assumir dois estados: *em operação* e *fora de serviço*. Estas unidades são interligadas segundo as características funcionais do sistema, compondo estruturas lógicas representadas na forma de diagramas em bloco. A partir dos diagramas, cálculos de probabilidade são utilizados para determinar os parâmetros de segurança de funcionamento do sistema em termos dos parâmetros associados às unidades. Dentre as estruturas possíveis, as mais usuais são as configurações série, paralelo e r-em-n. Adicionalmente, dependendo da natureza estrutural do sistema, os modelos combinatórios podem ser elaborados através de técnicas mais gerais como: grafos de confiabilidade, caminhos mínimos, métodos de decomposição, etc., [SHO 68], [BUZ 70a].

Estes modelos fornecem razoáveis estimativas dos parâmetros de segurança de funcionamento para sistemas relativamente simples. Todavia, mostram-se pouco adequados a representações detalhadas de alguns aspectos funcionais de sistemas mais sofisticados, como, por exemplo, reconfiguração em sistemas com redundância ativa.

A justificativa para tal limitação reside no fato de que estes modelos são baseados em hipóteses simplificadoras, como, por exemplo, comportamento estocasticamente independente do sistema. Na verdade, esta hipótese é bastante restritiva e, no caso de sistemas com redudância *standby*, a análise fica inviabilizada, uma vez que a taxa de falha das unidades varia em função delas se encontrarem operacionalmente na condição ativa ou reserva. Neste aspecto, as avaliações conduzem a resultados imprecisos.

Contrapondo-se às limitações apresentadas pelos dois primeiros tipos de modelos, os modelos de descrição de estados apresentam determinadas vantagens, principalmente por permitir a descrição do comportamento funcional dos sistemas através de distribuições

probabilísticas. Este aspecto é importante, uma vez que viabiliza, de forma simplificada, a representação de características particulares de funcionamento e, sobretudo, coaduna-se com a natureza estocástica dos sistemas redundantes e reparáveis.

Os modelos de descrição de estados possibilitam uma maneira sistemática de representar o funcionamento do sistema através de estados que descrevem o comportamento do mesmo e das taxas de transição entre os estados. Quando as taxas de transição são constantes, esta representação é descrita por um processo markoviano, constituindo, assim, os modelos de Markov. Estes modelos integram, segundo a descrição apresentada em [GEI 90], uma geração de ferramentas de avaliação que proporcionam uma maior precisão nos cálculos dos parâmetros de segurança de funcionamento, sendo amplamente indicados para sistemas com elevado nível de complexidade, onde se enquadram os equipamentos redundantes e reparáveis que incorporam uma parcela significativa dos sistemas de telecomunicações.

A modelagem segundo um processo markoviano permite avaliar o sistema em seus vários níveis de funcionamento. Inicialmente, nas fases de concepção e desenvolvimento, os modelos podem ser simples, representando apenas as tendências sistêmicas de segurança de funcionamento. À medida que o produto evolui, é possível optar por novas avaliações, onde modelos mais detalhados e complexos descrevem o comportamento funcional do sistema através de premissas mais realistas. Este procedimento é importante, uma vez que possibilita a obtenção de resultados mais precisos, muitas vezes necessários à homologação do produto, e à observação de alguns fatores que podem afetar a confiabilidade/disponibilidade dos sistemas tolerantes a defeitos.

Normalmente, os efeitos causados por estes fatores são difíceis de medir e de modelar. Entretanto, em função da relevância que alguns destes fatores assumem, a inclusão dos mesmos no modelo é muito mais uma condição indispensável do que uma oportunidade adicional de avaliação. Neste aspecto, a cobertura de defeitos pode ser considerada como um exemplo proeminente.

Dependendo da aplicação, diferentes objetivos de segurança de funcionamento devem ser perseguidos durante o desenvolvimento do sistema. Para tanto, são necessários diferentes métricas de análise, correspondendo aos parâmetros que melhor traduzem esses objetivos. Apesar desta diversidade métrica, a mesma concepção básica de modelos markovianos pode ser usada na estimativa dos parâmetros de segurança de funcionamento, observando, para cada caso, as particularidades de funcionamento inerentes ao requisito sob avaliação.

1.5 Medidas de Segurança de Funcionamento

As **medidas** ou **parâmetros** de segurança de funcionamento são definidos como probabilidade associadas a alguns atributos de segurança de funcionamento. A determinação destas medidas, através da avaliação do comportamento do sistema com respeito à ocorrência de

defeitos, tem por objetivo validar os mecanismos de tolerância a defeitos e, por extensão, verificar o atendimento do sistema às especificações de segurança de funcionamento.

Em geral, a vida de um sistema é percebida pelos usuários como uma alternância entre dois estados: operacional e fora de serviço. Quando no estado **operacional**, os serviços fornecidos pelo sistema atendem às especificações. No estado **fora de serviço**, o sistema não consegue fornecer os serviços de acordo com as especificações. Alguns sistemas, e.g., sistemas com partição de carga, admitem ainda um outro estado: serviço degradado. No estado caracterizado por **serviço degradado**, o sistema funciona com desempenho abaixo do esperado. A passagem de um estado operacional para o estado fora de serviço (ou para o estado de serviço degradado, no caso de alguns sistemas) é proporcionada pela ocorrência de falhas, ao passo que os reparos são responsáveis pela passagem inversa. Além disso, é possível a transição entre estados de operação no caso do sistema ser tolerante a defeitos.

Os parâmetros de segurança de funcionamento podem, então, ser obtidos a partir de características particulares associadas à alternância entre estes estados, tais como: a probabilidade de ocupação e o tempo de permanência nos estados. Neste aspecto, alguns parâmetros são definidos:

Confiabilidade: Uma medida da continuidade com que o serviço é fornecido de acordo com as especificações. Pode representar a probabilidade com que o sistema funcione corretamente por um dado intervalo de tempo, ou o tempo para que o sistema venha a falhar(ficar fora de serviço).

Disponibilidade: Uma medida que representa a probabilidade do sistema se encontrar num estado operacional num dado instante de tempo.

Mantenabilidade: Uma medida do tempo de reparo, e.g., tempo médio para reparo (MTTR), ou da continuidade com que o sistema pode se encontrar no estado fora de serviço.

Estes parâmetros são considerados como os mais usuais na quantificação dos atributos de segurança de funcionamento. Entretanto, é conveniente destacar mais alguns parâmetros (por exemplo, os tempos de permanência nos estados de um sistema), os quais, na verdade, correspondem a variáveis aleatórias associadas aos três primeiros. A seguir são apresentadas as definições destes parâmetros, juntamente com a nomenclatura usada na representação de suas médias estatísticas:

Tempo para falhar: Intervalo de tempo compreendido entre um instante em que o sistema está operacional (escolhido aleatoriamente) e o instante em que ocorre a próxima falha do sistema. É assumido que o intervalo de tempo entre o instante escolhido aleatoriamente e o instante de início de funcionamento do sistema é bastante longo. O valor médio deste parâmetro é representado por: MTTF.

Tempo para primeira falha: Intervalo de tempo compreendido entre o início de funcionamento do sistema e a ocorrência da primeira falha do sistema. O valor médio é representado por: MTFF.

Tempo entre falhas: Intervalo de tempo entre sucessivas falhas do sistema. Compreende a soma do tempo para falhar e o tempo para reparo. O valor médio é representado por: MTBF. É importante ressaltar que o MTBF tende para o MTTF quando o seu valor é muito maior que o MTTR, ou seja, uma vez que $MTBF = MTTF + MTTR$, quando $MTTF >> MTTR$, tem-se: $MTBF \simeq MTTF$.

Tempo em serviço: Intervalo de tempo compreendido entre um reparo do sistema e a próxima falha do sistema. O valor médio é denotado por MUT.

Tempo fora de serviço: Intervalo de tempo entre uma falha do sistema e o próximo reparo do sistema. O valor médio é denotado por MDT.

A escolha do parâmetro de segurança de funcionamento mais adequado à avaliação do sistema deve, sobretudo, ser baseada nas penalidades e custos decorrentes de uma falha. Neste aspecto, dois critérios merecem especial atenção: duração e frequência das falhas. No caso da duração da falha ser importante à aplicação do sistema, o parâmetro mais apropriado é a disponibilidade. No caso da frequência da falha ser importante, o tempo entre falhas ou o tempo para falhar são os parâmetros mais adequados.

Adicionalmente, o compromisso desempenho x confiabilidade tem representado um critério relevante, principalmente se consideradas as novas tendências das redes e sistemas de telecomunicações. O impacto deste compromisso na concepção dos projetos de sistema pode ser avaliada a partir das medidas de *performability* [MEY 78], [SMI 88], obtidas através da modelagem combinada de parâmetros de segurança de funcionamento e de desempenho.

1.6 Conclusões

O panorama até aqui esboçado permite notar a relevância da avaliação da segurança de funcionamento como suporte técnico-gerencial e instrumento de validação no processo de desenvolvimento de sistemas. Em termos gerais, este tipo de avaliação constitui um dos procedimentos que permitem garantir que os sistemas sejam desenvolvidos de modo a fornecer serviços de acordo com o nível de qualidade requerido. Como consequência, inúmeras técnicas e ferramentas têm sido elaboradas com a finalidade de fomentar tal tipo de avaliação.

No capítulo subsequente é apresentada uma ampla gama de técnicas e ferramentas utilizadas correntemente na avaliação dos parâmetros de segurança de funcionamento aqui apresentados, excetuando-se os parâmetros de *performability*, que, em função de seu caráter multidisciplinar, fogem ao escopo deste trabalho.

Capítulo 2

Segurança de Funcionamento: Técnicas de Avaliação

Neste capítulo são abordadas as principais técnicas de uso corrente na avaliação dos parâmetros de segurança de funcionamento. Uma vez que o objetivo é viabilizar a avaliação de sistemas de telecomunicações, do ponto de vista de falhas hardware[1] e sobretudo dos sistemas com aplicações comerciais, o conteúdo apresentado a seguir abrange os seguintes tópicos:

- Modelos de falha, relacionando a determinação dos parâmetros de segurança de funcionamento a partir da taxa de falha dos elementos do sistema.
- Modelos de descrição de estados, correspondendo a modelos de Markov com estado discreto e tempo contínuo, os quais possibilitam a avaliação dos parâmetros de segurança de funcionamento. Esta seção trata das propriedades gerais destes modelos, e dos procedimentos de elaboração e resolução dos mesmos, o que inclui a análise de características de funcionamento com influência marcante na segurança de funcionamento, e.g., cobertura e supervisão periódica de defeitos, e métodos de resolução para sistemas com grande número de estados, e.g., formulação matricial, agregação e randomização.
- Um panorama geral das ferramentas de avaliação, descrevendo, de forma sumária, as principais técnicas e utilitários computacionais existentes, os quais caracterizam o estado da arte das ferramentas de avaliação da segurança de funcionamento.

2.1 Modelos de falha

Esta seção apresenta a formulação de modelos de falha, como um primeiro passo na obtenção de meios de acesso aos parâmetros de segurança de funcionamento. Estes modelos estão relacionados com dados de taxa de falha dos elementos do sistema, através da teoria

[1] Métodos de avaliação da segurança de funcionamento com enfoque voltado para a integração das falhas hardware/software podem ser encontrados, por exemplo, em [LAP 84] e [STA 87]. Adicionalmente, métodos de avaliação direcionados para as falhas de concepção hardware, software e de sistema são encontrados em [MAR 91].

de probabilidade. Neste sentido são definidas algumas funções associadas ao processo de ocorrência de falhas, e.g., confiabilidade, MTTF, disponibilidade, funções distribuição e densidade de falha , e taxa de falha, com o propósito de introduzi-las neste trabalho e de evidenciar a relação existente entre elas.

2.1.1 Confiabilidade

A variável aleatória X é definida como o tempo de vida (ou o tempo para falhar) de um item, de forma que a probabilidade de falha em função do tempo é dada por $P(X < t) = F(t)$, onde $F(t)$ é a função de distribuição de falha . Assim, a probabilidade de que um item sobreviva até um determinado tempo pode ser definida como a confiabilidade do item, a qual é dada pela equação (2.1).

$$R(t) = P(X > t) = 1 - F(t) \quad . \tag{2.1}$$

Para uma população de N itens idênticos em funcionamento, é possível considerar que estes itens falham independentemente com probabilidade dada por $F(t)$. Consequentemente, após um intervalo de tempo t há $N_f(t)$ itens em falha e $N_s(t)$ itens sobreviventes, onde $N_f(t)$ e $N_s(t)$ são variáveis aleatórias e $N = N_f(t) + N_s(t)$. Representando $N_s(t)$ através de uma distribuição binomial com $p = R(t)$, o valor esperado de $N_s(t)$, designado por $n(t)$ é, como pode ser observado em [SHO 68], dado por:

$$n(t) \triangleq E[N_s(t)] = NR(t) \quad , \tag{2.2}$$

ou seja,

$$R(t) = \frac{n(t)}{N} \quad . \tag{2.3}$$

Substituindo (2.3) em (2.1), é possível escrever:

$$F(t) = 1 - \frac{n(t)}{N} \quad . \tag{2.4}$$

Como a função densidade de probabilidade $f(t)$ é expressa por $f(t) = \frac{dF(t)}{dt}$, tem-se:

$$f(t) = -\frac{1}{N}\frac{dn(t)}{dt} \equiv \frac{1}{N}\lim_{\Delta t \to 0}\frac{n(t) - n(t + \Delta t)}{\Delta t} \quad . \tag{2.5}$$

Dividindo (2.5) por $n(t)$:

$$\frac{f(t)N}{n(t)} = \lim_{\Delta t \to 0}\frac{n(t) - n(t + \Delta t)}{n(t)\Delta t} \quad . \tag{2.6}$$

O termo do lado direito da equação (2.6) é a definição da taxa de falha $\lambda(t)$, a qual está relacionada com a função densidade de falha normalizada em termos de $n(t)$. Assim,

$$\lambda(t) = \frac{f(t)N}{n(t)} = \frac{f(t)N}{R(t)N} = \frac{f(t)}{R(t)} \quad . \tag{2.7}$$

É importante notar que em se tratando de distribuição exponencial, a taxa de falha é caracterizada por um valor constante, como pode ser notado através da relação apresentada pela equação (2.7), ou seja,

$$\lambda(t) = \frac{f(t)}{R(t)} = \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda \quad . \tag{2.8}$$

Em outras palavras, $\lambda(t)\Delta t$ representa a probabildade de que um item falhe no intervalo $(t, t + \Delta t]$, dado que ele se encontre operacional no instante t.

A relação entre confiabilidade e taxa de falha também pode ser obtida a partir de (2.7). Utilizando-se a equação (2.1) e a condição de que $f(t) = \frac{dF(t)}{dt}$, é possível escrever:

$$\lambda(t) = \frac{f(t)}{R(t)} = \frac{-dR(t)}{R(t)dt} \quad . \tag{2.9}$$

Integrando ambos os lados de (2.9) no intervalo [0,t], e usando a condição de contorno $R(0) = 1$, tem-se:

$$\int_0^t \lambda(x)dx = -\ln R(t)$$

logo,

$$R(t) = e^{-\int_0^t \lambda(x)dx} \quad . \tag{2.10}$$

2.1.2 MTTF

Um outro aspecto importante consiste na quantificação do tempo de funcionamento de um item. Para tanto, são normalmente usados o tempo médio para falhar (MTTF) ou o tempo médio entre falhas (MTBF). Frequentemente estes dois parâmetros são utilizados como equivalentes, entretanto, esta aproximação só ocorre quando o MTTR é muito menor que o MTTF.

Uma maneira de quantificar o tempo de funcionamento de um item pode ser obtida com a utilização de seu valor médio, que, por sua vez, define o MTTF, ou seja:

$$MTTF = E[t] = \int_0^\infty t f(t)dt \quad .$$

Escrevendo de outra forma, tem-se:

$$MTTF = -\int_0^\infty t\frac{dR(t)dt}{dt} = -\int_0^\infty t dR(t) \quad . \tag{2.11}$$

A resolução de (2.11) leva à seguinte expressão:

$$MTTF = -tR(t)|_0^\infty + \int_0^\infty R(t)dt \quad . \tag{2.12}$$

Examinando os limites superior e inferior de $tR(t)$, podemos constatar que ambos tendem a zero, desde que $R(0) = 1$. Desta forma,

$$MTTF = \int_0^\infty R(t)dt \quad . \tag{2.13}$$

Para sistemas com arquiteturas complexas, a obtenção destes parâmetros está atrelada a análises que decompõem o sistema em unidades ou subsistemas, interligando-as segundo os princípios de funcionamento do sistema. A formulação destes modelos obedece a esta estrutura lógica, e a confiabilidade do sistema é determinada em termos da confiabilidade associada às subdivisões em que o sistema foi decomposto.

Os modelos decorrentes deste tipo de abordagem, normalmente conhecidos como estruturais, cujos mais usuais estão associados a configurações série, paralelo e r-em-m, [SHO 68], [RAC], [TRI 82], são baseados na hipótese de independência entre as unidades, quanto à ocorrência de falhas e reparações.

2.1.3 Disponibilidade

Mantendo a hipótese de independência para os modelos estruturais de sistemas reparáveis, ou seja, assumindo que o processo de reparação ocorre como se cada unidade tivesse seu próprio reparador[2], é possível obter a expressão de indisponibilidade através da abordagem apresentada em [SOU 87]. Esta abordagem considera que cada unidade tem dois estados possíveis:

- **Estado operacional**: no qual a unidade ou o sistema funciona corretamente até a ocorrência de uma falha. A taxa de falha é constante e representada por λ, de forma que a probabilidade de uma unidade falhar no intervalo $(t, t+\Delta t]$, dado que no instante t ela está no estado operacional, é $\lambda\Delta t$.
- **Estado de falha**: no qual a unidade ou o sistema não apresenta condições de funcionamento correto em função da ocorrência de uma falha. Esta situação de não operacionalidade dura até a reparação da falha. A taxa de reparo é constante e dada por μ, de forma que a probabilidade de uma unidade ser reparada no intervalo $(t, t+\Delta t]$, dado que no instante t ela se encontra no estado de falha, é $\mu\Delta t$.

Sendo $P_o(t)$ e $P_f(t)$ as probabilidades de ocupação dos estados operacional e fora de serviço, respectivamente, as equações que descrevem a mudança entre estes estados são dadas por (2.14) e (2.15).

$$P_o(t+\Delta t) = P_o(t)(1-\lambda\Delta t) + P_f(t)\mu\Delta t \tag{2.14}$$

$$P_f(t+\Delta t) = P_o(t)\lambda\Delta t + P_f(1-\mu\Delta t) \tag{2.15}$$

onde $(1-\lambda\Delta t)$ representa a probabilidade de não ocorrência de falha no período Δt, e $(1-\mu\Delta t)$ a probabilidade de que não haja reparação no período Δt.

[2]Em se tratando de sistemas reparáveis sem redundancia, esta é uma hipótese razoável, uma vez que cada sistema pode ter seu próprio reparador

Rearranjando as equações (2.14)e (2.15), e tomando o limite para Δt se aproximando de zero, obtem-se as seguintes equações diferenciais:

$$\frac{dP_o(t)}{dt} = -P_o(t)\lambda + P_f(t)\mu \tag{2.16}$$

$$\frac{dP_f(t)}{dt} = P_o(t)\lambda - P_f(t)\mu \ . \tag{2.17}$$

Assumindo as condições iniciais $P_o(0) = 1$ e $P_f(0) = 0$ e utilizando a transformada de Laplace nas equações (2.16) e (2.17), onde $g^*(s)$ representa a transformada de Laplace de uma função $g(t)$, tem-se:

$$sP_o^*(s) - 1 = -P_o^*(s)\lambda + P_f^*(s)\mu \tag{2.18}$$

$$sP_f^*(s) = P_o^*(s)\lambda - P_f^*(s)\mu \ . \tag{2.19}$$

Resolvendo este sistema de equações, obtem-se a expressão (2.20) de $P_o^*(s)$:

$$P_o^*(s) = \frac{(s+\mu)}{[(s+\lambda+\mu)s]} \ . \tag{2.20}$$

Invertendo para o domínio do tempo, chega-se à expressão (2.21):

$$P_o(t) = \frac{\mu}{(\mu+\lambda)} + \frac{\lambda}{(\lambda+\mu)} e^{-(\lambda+\mu)t} \ , \tag{2.21}$$

a qual corresponde à expressão de disponibilidade da unidade/sistema.

Para os sistemas eletrônicos, a duração da missão t é muito maior que$(\lambda+\mu)^{-1}$, o que torna praticamente nulo o segundo termo da equação (2.21). Desta forma, é possível expressar a disponibilidade através de um fator constante, considerando o comportamento do sistema (no que tange ao processo de falha-reparação) em regime estacionário. Consequentemente, a disponibilidade D é dada pela equação (2.22).

$$D = \frac{\mu}{\mu+\lambda} \ . \tag{2.22}$$

A hipótese de independência entre as unidades redundantes, no que diz respeito aos processos de falha e reparação, impõe um caráter restritivo à análise, e não pode ser adotada nos casos de redundância ativa, quando a taxa de falha da unidade varia em função da unidade se encontrar na condição ativa ou reserva. Além disso, a hipótese de independência também não pode ser usada no caso de sistemas reparáveis com número limitado de reparadores, posto que existe concorrência entre os processos de falha e reparação (i.e., durante a reparação existe a probabilidade de ocorrência de outras falhas).

Neste aspecto, os modelos de descrição de estados apresentam-se como ferramentas mais adequadas à avaliação de sistemas reparáveis e com características de funcionamento complexas, uma vez que admitem a hipótese de dependência entre os processos de falha e reparação.

2.2 Modelo de Descrição de Estados

A modelagem através de estados que descrevam o comportamento operacional de um sistema torna-se adequada aos casos de sistemas que possuem redundância em *standby* e estão sujeitos a reparos e falhas dependentes. No caso específico dos equipamentos eletrônicos, cujas taxas de falha e de reparo podem ser consideradas constantes, a modelagem através de processos markovianos constitui uma poderosa ferramenta de acesso às medidas de segurança de funcionamento.

2.2.1 Propriedade dos modelos de Markov

Os modelos de Markov são caracterizados por duas variáveis aleatórias (v.a.): a v.a. que representa o estado do sistema, X, e a v.a. que representa o tempo de observação, t. Consequentemente, surgem quatro tipos de modelo, uma vez que as variáveis X e t tanto podem ser discretas como contínuas. Dentre estes tipos de modelo, um, em particular, assume um papel de grande importância na avaliação da segurança de funcionamento: aqueles que descrevem o sistema como um processo markoviano com estado discreto e tempo contínuo.

Modelos com estado discreto e tempo discreto [KLE 75], bem como modelos semi-markovianos [BAR 65], [HOW 71], podem ser utilizados, porém numa frequência bem menor. Os modelos com tempo discreto dificilmente são utilizados, salvo situações muito específicas em que o tempo de permanência nos estados é melhor representado através de uma distribuição geométrica. Já os modelos semi-markovianos são adotados diante da necessidade de representar o tempo de permanência nos estados por meio de uma distribuição diferente das distribuições sem memória, o que pode ocorrer em se tratando de sistemas sujeitos a fatores externos que variam com o tempo da missão, e que requerem avaliações com resultados precisos. O apêndice A ilustra a avaliação de indisponibilidade de um sistema redundante a partir de um processo semi-markoviano.

Além do caráter particular em termos da aplicação destes dois tipos de modelos, os sistemas que, por ventura, requeiram a utilização de um dos dois não constituem parcela significativa entre os sistemas de telecomunicações voltados a aplicações comerciais. Desta forma, nesta subsecção são abordados apenas os modelos com estado discreto e tempo contínuo.

De uma forma geral, um processo estocástico $X(t)$ é chamado um processo markoviano se, para qualquer $t_o < t_1 < t_2 < ... < t_n < t$, a distribuição condicional de $X(t)$ para dados valores de $X(t_0), X(t_i), ..., X(t_n)$ depende apenas do valor de $X(t_n)$, ou seja:

$$P[X(t) \leq x | X(t_n) = x_n, X(t_{n-1}) = x_{n-1}, ..., X(t_0) = x_0] = \\ = P[X(t) \leq x | X(t_n) = x_n] \quad . \tag{2.23}$$

Um modelo descrito por um processo de Markov é definido por um conjunto de probabilidades p_{ij}, as quais representam a probabilidade de transição de um estado i para um estado j. Uma característica importante destes modelos reside no fato de que as probabilidades de transição p_{ij} dependem apenas dos estados i e j, sendo independente dos demais

estados.

Desta maneira, num processo markoviano, os acontecimentos passados são todos resumidos no estado corrente, de forma que a distribuição para o tempo Y de ocupação de um dado estado deve ser sem memória, ou seja:

$$P[Y \leq t + t_0 | Y \geq t_0] = P[Y \leq t] \quad . \tag{2.24}$$

Nos modelos de Markov, esta propriedade pode ser percebida através da probabilidade de transição entre dois estados. Sendo X_{ij} a variável aleatória que representa o tempo de transição do estado i para o estado j, e sabendo-se que no instante t o processo markoviano está no estado i, é possível escrever que a probabilidade $p_{ij}(t)\Delta t$ de que no intervalo $(t, t + \Delta t]$ o processo esteja no estado j, é dada por:

$$p_{ij}(t)\Delta t = P[t < X_{ij} \leq t + \Delta t | X_{ij} > t] \quad . \tag{2.25}$$

Utilizando a equação (2.7) é possível expressar $p_{ij}(t)$ da seguinte maneira:

$$p_{ij}(t) = \frac{f_{ij}(t)}{R_{ij}(t)} \quad , \tag{2.26}$$

onde $f_{ij}(t)$ e $R_{ij}(t)$ representam, respectivamente, as funções densidade de probabilidade e de sobrevivência (ou confiabilidade) da variável aleatória X_{ij}.

No caso de X_{ij} ser exponencialmente distribuída, pode-se verificar através de (2.26) que $p_{ij}(t)$ é constante e igual a λ_{ij}, onde λ_{ij} é a taxa de falha da unidade ativa do sistema. Nestes casos, o modelo é chamado de homogêneo.

Desta forma, é possível observar que, nos modelos de Markov, a propriedade do tempo de permanência nos estados ter distribuição sem memória se adapta perfeitamente à condição de que as taxas de falha dos sistemas eletrônicos podem ser consideradas constantes. Esta hipótese de taxas de falha constantes pode ser aplicada aos sistemas de telecomunicações com aplicações comerciais, sem prejuízo de precisão para os resultados obtidos, uma vez que para estes sistemas o principal interesse se concentra na determinação das probabilidades estacionárias de ocupação dos estados.

Este aspecto deixa evidente a aplicabilidade dos modelos de Markov na avaliação dos parâmetros de segurança de funcionamento, principalmente se for considerado o fato de que estes modelos proporcionam descrições mais realistas em termos do comportamento operacional dos sistemas, e podem ser resolvidos através de técnicas razoavelmente conhecidas.

Com relação à terminologia associada aos modelos de markov, é importante definir os seguintes termos: processo ergódico, e estados recorrente, transiente e absorvedor.

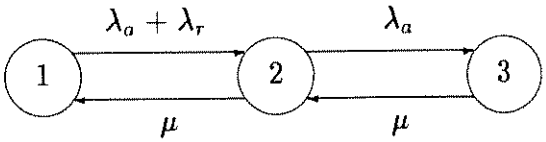
- **Estado recorrente**: um estado é considerado recorrente se o retorno a este estado ocorre com probabilidade igual a um.
- **Estado transiente**: um estado é dito transiente se o retorno a este estado é um evento não certo.
- **Estado absorvedor**: é a designação utilizada a um estado que, uma vez alcançado, a probabilidade de que o processo permaneça neste estado é 1. Desta forma, a probabilidade do processo se encontrar num estado absorvedor, quando $t \rightarrow \infty$, é 1.
- **Processo ergódico**: um processo é caracterizado como ergódico quando não possui estados absorvedores, sendo constituído apenas por estados ergódicos (ou recorrentes não nulos). Um estado ergódico é um estado recorrente com tempo médio de recorrência finito. Num processo ergódico, a probabilidade estacionária de ocupação de qualquer estado independe do estado inicial ($t = 0$) do processo.

Na elaboração dos modelos de Markov, o interesse está na determinação da probabilidade $P_i(t)$ de que o processo esteja no estado i no instante t, sendo necessário, para tanto, formular as equações de transição de estado. Com o intuito de exemplificar a formulação destas equações, tomemos como base o modelo da figura 2.1. Este modelo corresponde à descrição de um sistema redundante sujeito à reparação, onde λ_a e λ_r representam a taxa de falha da unidade ativa e reserva, respectivamente, e μ é a taxa de reparo. A partir deste modelo é possível escrever as equações que regem a probabilidade de transição entre os estados, conforme mostrado a seguir:

$$P_1(t+dt) = [1-(\lambda_a+\lambda_r)dt]P_1(t) + \mu dt P_2(t)$$
$$P_2(t+dt) = (\lambda_a+\lambda_r)dt P_1(t) + [1-(\lambda_a+\mu)dt]P_2(t) + \mu dt P_3(t)$$
$$P_3(t+dt) = \lambda_a dt P_2(t) + [1-\mu dt]P_3(t) \quad .$$

No limite, quando $dt \rightarrow 0$, tem-se:

$$\frac{dP_1(t)}{dt} = -(\lambda_a+\lambda_r)P_1(t) + \mu P_2(t)$$
$$\frac{dP_2(t)}{dt} = (\lambda_a+\lambda_r)P_1(t) - (\lambda_a+\mu)P_2(t) + \mu P_3(t)$$
$$\frac{dP_3(t)}{dt} = \lambda_a P_2(t) - \mu P_3(t)$$



Estado 1 → As duas unidades se encontram em condições de operação

Estado 2 → Uma unidade está em falha e a outra em operação

Estado 3 → As duas unidades estão em falha e, consequentemente, o sistema não consegue fornecer o serviço.

Figura 2.1 - Modelo de Markov para sistema redundante 1:1 com reparação

Assim, a obtenção das probabilidades $P_i(t), i = 1, 2, ..., N$ num processo markoviano com N estados é feita a partir da solução de (N-1) equações diferenciais, obtidas do modelo, mais a condição $\sum_{i=1}^{N} P_i(t) = 1$. Conhecendo-se as probabilidades $P_i(t)$, o parâmetro de segurança de funcionamento que se deseja avaliar pode ser obtido. Por exemplo, no modelo da figura 2.1, a probabilidade de ocupação dos estados 1 e 2 corresponde à disponibilidade do sistema $D(t)$, logo:

$$D(t) = P_1(t) + P_2(t) \ .$$

Adicionalmente, alterando o modelo da figura 2.1 de forma a considerar o estado 3 como absorvedor, o que é feito removendo-se a taxa de transição μ entre os estados 3 e 2, é possível obter a confiabilidade deste sistema a partir das probabilidades P_1 e P_2 do modelo alterado, ou seja,

$$R(t) = P_1(t) + P_2(t) \ .$$

Além disso, quando $t \rightarrow \infty$, a confiabilidade converge para o $MTTF$, consequentemente, $MTTF = R(\infty)$. Em termos da transformada de Laplace, tem-se:

$$MTTF = R^*(s)|_{s=0} \ .$$

Na maior parte das vezes, o interesse da avaliação se concentra na probabilidade estacionária de ocupação dos estados, o que facilita a resolução do sistema de equações, uma vez que em regime estacionário $\frac{dP(t)}{dt} = 0$.

Em termos gerais, no caso de modelos com número reduzido de estados, o acesso aos parâmetros de segurança de funcionamento é facilmente obtido. Entretanto, para modelos de sistemas complexos, com um elevado número de estados, a obtenção das expressões dos parâmetros de segurança de funcionamento não é uma tarefa simples, o que torna indispensável o emprego de técnicas matriciais e programas computacionais na avaliação destes parâmetros.

2.2.2 Resolução matricial

A utilização de técnicas matriciais na obtenção das expressões dos parâmetros de segurança de funcionamento propicia a formulação de programas computacionais, constituindo um procedimento amplamente adequado à avaliação de sistemas complexos. A resolução matricial descrita aqui é baseada na abordagem apresentada em [SOU 87].

Seja um processo markoviano com N estados e taxa de transição do estado i para o estado j, λ_{ij}, constante para todos os i, j. Seja $P_i(t)$ a probabilidade de ocupação do estado i no instante t, e Λ a matriz de taxas de transição com elementos λ_{ij}, para $i \neq j$ e

$$\lambda_{ii} = - \sum_{k=1, k \neq i}^{N} \lambda_{ik} \quad .$$

Escrevendo as equações de transição de estado, tem-se:

$$P_i(t + dt) = \sum_{k=1}^{N} P_k(t) \lambda_{ki} dt + P_i(t)$$

com $\lambda_{ii} = - \sum_{k=1, k \neq i}^{N} \lambda_{ik}$. No limite, quando $dt \rightarrow 0$

$$\frac{dP_i(t)}{dt} = \dot{P}_i(t) = \sum_{k=1}^{N} P_k(t) \lambda_{ki} \quad . \tag{2.27}$$

Seja $\vec{P}(t)$ o vetor das probabilidades de ocupação dos estados, ou seja, $\vec{P}(t) = [P_1(t) ... P_N(t)]$. Assim, é possível escrever (2.27) na forma matricial dada a seguir:

$$\dot{\vec{P}}(t) = \vec{P}(t) \Lambda \quad , \tag{2.28}$$

onde Λ é a matriz de taxas de transição com elementos λ_{ij}.

Uma vez que o interesse nos modelos de acesso às grandezas de segurança de funcionamento gira em torno da probabilidade que o sistema esteja num estado operacional ou de falha, ou do tempo médio em que o sistema passa nesses estados, é conveniente dividir os estados do modelo em dois conjuntos: um dos estados operacionais (o), e outro dos estados de falha (f), nos quais o sistema está fora de serviço.

Baseado nesta divisão, as transições podem ser separadas segundo a classificação dos estados de origem e de destino em um destes dois estados, de tal forma que a matriz Λ pode ser escrita da seguinte forma:

$$\Lambda = \begin{bmatrix} \Lambda_{oo} & \Lambda_{of} \\ \Lambda_{fo} & \Lambda_{ff} \end{bmatrix}$$

onde Λ_{oo}, por exemplo, representa a submatriz de taxas de transição entre estados operacionais.

Procedimento similar pode ser aplicado ao vetor $\vec{P}(t)$, de tal sorte que

$$\vec{P}(t) = [\vec{P}_o(t) \quad \vec{P}_f(t)] \ .$$

Com isto, as equações de transição de estado podem ser escritas na forma:

$$[\dot{\vec{P}}_o(t) \quad \dot{\vec{P}}_f(t)] = [\vec{P}_o(t) \quad \vec{P}_f(t)] \begin{bmatrix} \Lambda_{oo} & \Lambda_{of} \\ \Lambda_{fo} & \Lambda_{ff} \end{bmatrix} \ . \tag{2.29}$$

Como exemplificações da formulação matricial, a seguir são apresentados os processos de obtenção de alguns parâmetros de segurança de funcionamento, a saber: confiabilidade, MTFF, MTTF e disponibilidade.

Expressão da confiabilidade/MTTF

Na avaliação da confiabilidade, e por conseguinte do MTTF, os estados de falha são considerados como absorvedores. Portanto, na matriz de taxas de transição, Λ_{fo} e Λ_{ff} só possuem elementos nulos. Deste modo, a equação matricial (2.29) pode ser expandida da seguinte forma:

$$\dot{\vec{P}}_o(t) = \vec{P}_o(t)\Lambda_{oo} \tag{2.30}$$

$$\dot{\vec{P}}_f(t) = \vec{P}_f(t)\Lambda_{of} \ . \tag{2.31}$$

Fazendo a transformada de Laplace da equação (2.30), e sabendo-se que $R(t)$ é dada por $R(t) = \vec{P}_o(t)1_k$, onde 1_k é o vetor coluna de k elementos unitários, com k igual ao número de estados operacionais, e I é a matriz identidade,tem-se:

$$R^*(s) = \vec{P}_o^*(s)1_k = \vec{P}_o(0)[sI - \Lambda_{oo}]^{-1}1_k \ . \tag{2.32}$$

Como o $MTTF = R^*(s)|_{s=0}$, segue-se:

$$MTTF = \vec{P}_o^*(s)|_{s=0} = \vec{P}_o(0)[-\Lambda_{oo}]^{-1}1_k \ , \tag{2.33}$$

onde $\vec{P}_o(0)$ é o vetor das probabilidades iniciais de ocupação dos estados operacionais.

Expressão do MTFF

Na abordagem apresentada em [BUZ 70b], a expressão que determina o tempo médio para primeira falha é obtida da seguinte maneira:

Sendo o estado 1 o estado no qual todos os componentes estão em perfeitas condições de operação, então $\vec{P}_o(0) = [1 \quad 0 \ ... \ 0]$, e sendo $\vec{T} \equiv [T_1 \quad T_2 \ ... \ T_k]$ o vetor dos tempos de permanência nos estados operacionais, o $MTFF$ pode ser obtido através da seguinte

expressão:

$$MTFF = \sum_{i=1}^{k} T_i \quad , \tag{2.34}$$

onde k é o número de estados operacionais e T_i é obtido resolvendo-se o seguinte sistema de k equações:

$$-\vec{T}\Lambda_{oo} \equiv [1 \quad 0 \ldots 0] \quad ,$$

ondeΛ_{oo} é a submatriz $k \times k$ de taxas de transição.

Expressão do MTTF

Conhecendo-se as probabilidades estacionárias de ocupação dos estados, representados pelo vetor $\vec{P} = [P_o \quad P_f]$, o $MTTF$ também pode ser obtido através da abordagem apresentada em [BUZ 70b], ou seja,

$$MTTF = \frac{\vec{P}_o(-\Lambda_{oo})^{-1}1_k}{\vec{P}_o 1_k} \quad , \tag{2.35}$$

onde 1_k é o vetor coluna de k elementos unitários, k é o número de estados operacionais e Λ_{oo} é a submatriz $k \times k$ de taxas de transição.

Expressão da disponibilidade

Na avaliação de disponibilidade, é considerado um processo markoviano ergódico e o interesse reside na determinação das probabilidades de ocupação dos estados. Neste caso, as submatrizes Λ_{fo} e Λ_{ff} apresentam elementos não nulos.

Partindo da equação diferencial para N estados,

$$\dot{\vec{P}}(t) = \vec{P}(t)\Lambda \quad ,$$

e sabendo-se que em regime estacionário $\dot{\vec{P}}(t) = 0$ e que $\vec{P} = [P_i]$ é definido como o vetor de probabilidade estacionária, pode-se escrever:

$$\vec{P}\Lambda = 0 \quad .$$

Utilizando a condição $\sum_{i=1}^{N} P_i = 1$, na forma matricial, $\vec{P}1_N = 1$, onde 1_N é o vetor coluna de N elementos unitários, e substituindo a primeira coluna de Λ por 1, resultando na matriz Λ_m, tem-se:

$$\vec{P}\Lambda_m = [1 \quad 0 \ldots 0] \quad ,$$

logo,

$$\vec{P} = [1 \quad 0 \ldots 0][\Lambda_m]^{-1} \quad .$$

Como a disponibilidade D é dada por $D = \vec{P}1_k$, onde 1_k é o vetor coluna de k elementos unitários, com k igual ao número de estados operacionais do modelo, é possível escrever:

$$D = \vec{P}1_k = [1 \quad 0 \ldots 0][\Lambda_m]^{-1}1_k \quad . \tag{2.36}$$

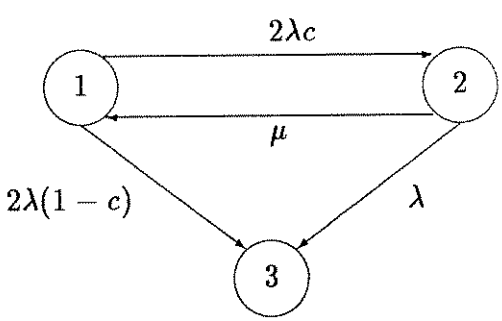
O procedimento de determinação de algumas destas expressões é apresentado a seguir, a título ilustrativo, através de uma avaliação que inclui a cobertura de defeitos. A modelagem envolvendo o fator de cobertura proporciona avaliações mais realistas, uma vez que este fator tem uma influência marcante nas grandezas de segurança de funcionamento.

2.2.3 Influência da cobertura

Conforme já mencionado no capítulo 1, a utilização efetiva da redundância está atrelada ao funcionamento correto dos mecanismos de supervisão. A probabilidade que estes mecanismos funcionem corretamente dado a ocorrência de falha de um item do sistema redundante é representada pelo fator de cobertura (c). Num sentido mais quantitativo, este parâmetro corresponde à proporção de defeitos que o sistema tem condições de detectar e localizar, além de garantir a continuação do serviço a partir do acionamento dos itens reservas, ou seja:

$$c = P\{\text{ detecção, localização e reconfiguração } | \text{ ocorrência de uma falha}\}$$

Com o propósito de avaliar o impacto da cobertura nos parâmetros de segurança de funcionamento, a seguir é considerado o modelo de um sistema com duas unidades iguais e redundantes, funcionando segundo o princípio de redundância *hot standby*. Neste tipo de redundância, uma unidade é colocada na condição ativa e a outra na condição reserva, a qual se encontra apta a passar à condição ativa mediante a ocorrência de falha da primeira.



Estado 1 → as duas unidades estão funcionando normalmente.

Estado 2 → uma unidade funcionando corretamente após falha coberta da outra unidade.

Estado 3 → as duas unidades estão em falha, ou falha não coberta da unidade ativa. Representa um estado de falha.

Figura 2.2 - Modelo de Markov para um sistema redundante *hot standby* 1:1.

Baseado nestas características de funcionamento, é possível descrever o comportamento operacional do sistema conforme apresentado no modelo da figura 2.2. Neste modelo, λ representa a taxa de falha de cada unidade, μ é a taxa de reparo, e c o fator de cobertura, sendo adotadas as seguintes hipóteses:

i) A cobertura é considerada instantânea e relacionada com defeitos hardware permanentes.

ii) É assumido que após um reparo, o sistema volta às condições normais de operação, e que há apenas um reparador com taxa μ.

Adotando a formulação matricial no sentido de determinar a expressão da confiabilidade e do $MTTF$, tem-se:

$$\Lambda = \begin{bmatrix} -2\lambda & 2\lambda c & 2\lambda(1-c) \\ \mu & -(\mu+\lambda) & \lambda \\ 0 & 0 & 0 \end{bmatrix}$$

Utilizando a expressão (2.33) para $P_o(0) = [1 \quad 0]$, obtem-se:

$$MTTF = [1 \quad 0] \begin{bmatrix} 2\lambda & -2\lambda c \\ -\mu & (\mu+\lambda) \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad .$$

logo,

$$MTTF = \frac{\lambda + \mu + 2\lambda c}{2\lambda^2 + (1-c)2\lambda\mu} = \frac{1 + (\lambda/\mu)(1+2c)}{(2\lambda^2/\mu + [1 + (1-c)(\mu/\lambda)]} \quad . \tag{2.37}$$

Em se tratando de sistemas eletrônicos, $\mu \gg \lambda$, consequentemente, é possível aproximar a equação (2.37) para

$$MTTF = \frac{1}{(2\lambda^2/\mu)[1 + (1-c)(\mu/\lambda)]} \quad . \tag{2.38}$$

Dividindo o caso ideal, onde $c = 1$ ($MTTF_{c=1}$), pela expressão (2.38), ou seja,

$$\frac{MTTF_{c=1}}{MTTF} = 1 + (1-c)\frac{\mu}{\lambda} \quad , \tag{2.39}$$

pode-se notar que para coberturas imperfeitas ($c < 1$), o $MTTF$ é reduzido pelo fator $1 + (1-c)\frac{\mu}{\lambda}$. Para aferir melhor esta influência, adotemos um exemplo baseado em valores típicos. Supondo que uma unidade é acometida por falha a cada 6 meses ($\lambda = 2.3 \times 10^{-4}$) e que ela pode ser reparada em $2h$ ($\mu = 0.5$), observa-se que o MTTF para c=0.98 é aproximadamente 44 vezes menor que o $MTTF$ para o caso ideal ($c = 1$).

Para avaliar o impacto da cobertura na disponibilidade (ou indisponibilidade) de um sistema, adotemos o modelo da figura 2.3. Este modelo corresponde a um sistema redundante idêntico ao descrito na figura 2.2, com a diferença de que agora o sistema tem coberturas diferenciadas para as unidades ativa e reserva, e está sujeito a reconfigurações manuais (com taxa μ') quando da ocorrência de falha não coberta da unidade ativa.

O modelo da figura 2.3 leva em consideração as mesmas hipóteses adotadas para o modelo anterior, e possui a seguinte notação:

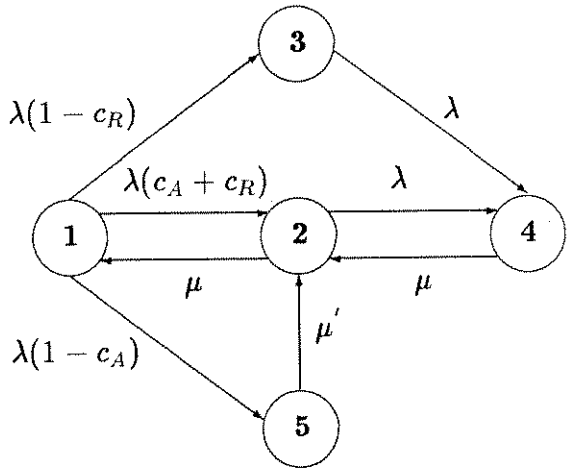
$\lambda \rightarrow$ taxa de falha hardware relativa a cada unidade.

$c_A \rightarrow$ fator de cobertura da unidade ativa.

$c_R \rightarrow$ fator de cobertura da unidade reserva.

$\mu \rightarrow$ taxa de reparo (falha permanente), $\mu = 1/MTTR$.

$\mu' \rightarrow$ taxa de reparo (falha permanente) referente a um procedimento de reconfiguração manual, $\mu' > \mu$.



Estado 1 $\rightarrow$ duas unidades em condições operacionais.

Estado 2 $\rightarrow$ uma unidade com falha coberta e a outra em condições operacionais.

Estado 3 $\rightarrow$ falha não coberta da unidade reserva. Estado operacional.

Estado 4 $\rightarrow$ falha das duas unidades. Estado não operacional.

Estado 5 $\rightarrow$ falha não coberta da unidade ativa. Estado não operacional.

Figura 2.3 - Modelo de Markov para um sistema redundante (*hot standby 1:1*) com coberturas diferenciadas.

É importante ressaltar que a cobertura associada à unidade reserva representa a capacidade do sistema em detectar e localizar/sinalizar os defeitos que possam levar esta unidade à condição de falha. Esta cobertura não abrange a capacidade de reconfiguração, uma vez que a unidade já se encontra na reserva.

Resolvendo o sistema de equações de estado para regime estacionário, o qual inclui a condição $\sum_i P_i = 1$, onde P_i é a probabilidade estacionária de ocupação do estado i, obtem-se a expressão de indisponibilidade I. A expressão de I é apresentada em (2.40) e corresponde à probabilidade estacionária de ocupação dos estados 4 e 5.

$$I = \frac{(1 - C_A)\lambda/\mu' + 2\lambda^2/\mu^2 + (1 - c_R)\lambda/\mu}{1 + 2\lambda/\mu + 2\lambda^2/\mu^2 + (1 - c_R)(1 + \lambda/\mu) + (1 - c_A)\lambda/\mu'} \tag{2.40}$$

A tabela 2.1 apresenta alguns valores de I em função da cobertura. Estes valores correspondem ao caso em que $c_A = c_R = c$, $\mu = 1$, $\mu' = 4$ e $\lambda = 10^{-4} falhas/h$, e mostram que a indisponibilidade tem seu valor diminuído em mais de 140 vezes, quando a cobertura passa de 0.9 a 1.

I	c
1.14×10^{-5}	0.9
1.26×10^{-6}	0.99
1.45×10^{-7}	0.999
8×10^{-8}	1

Tabela 2.1 - Indisponibilidade de um sistema redundante (*hot standby 1:1*) X cobertura.

As curvas apresentadas na figura 2.4 ilustram a influência da cobertura na indisponibilidade do sistema descrito pelo modelo da figura 2.3. Estas curvas foram obtidas considerando-se as mesmas taxas, de reparo e de falha, mencionadas para a tabela 2.1. Numa das curvas, o c_A é mantido constante, enquanto o c_R assume valores de 0.9 a 1. Na outra curva, a situação se inverte. A finalidade destas curvas é mostrar que, em função das características operacionais do sistema, a cobertura associada à unidade reserva exerce uma influência maior sobre a indisponibilidade do sistema.

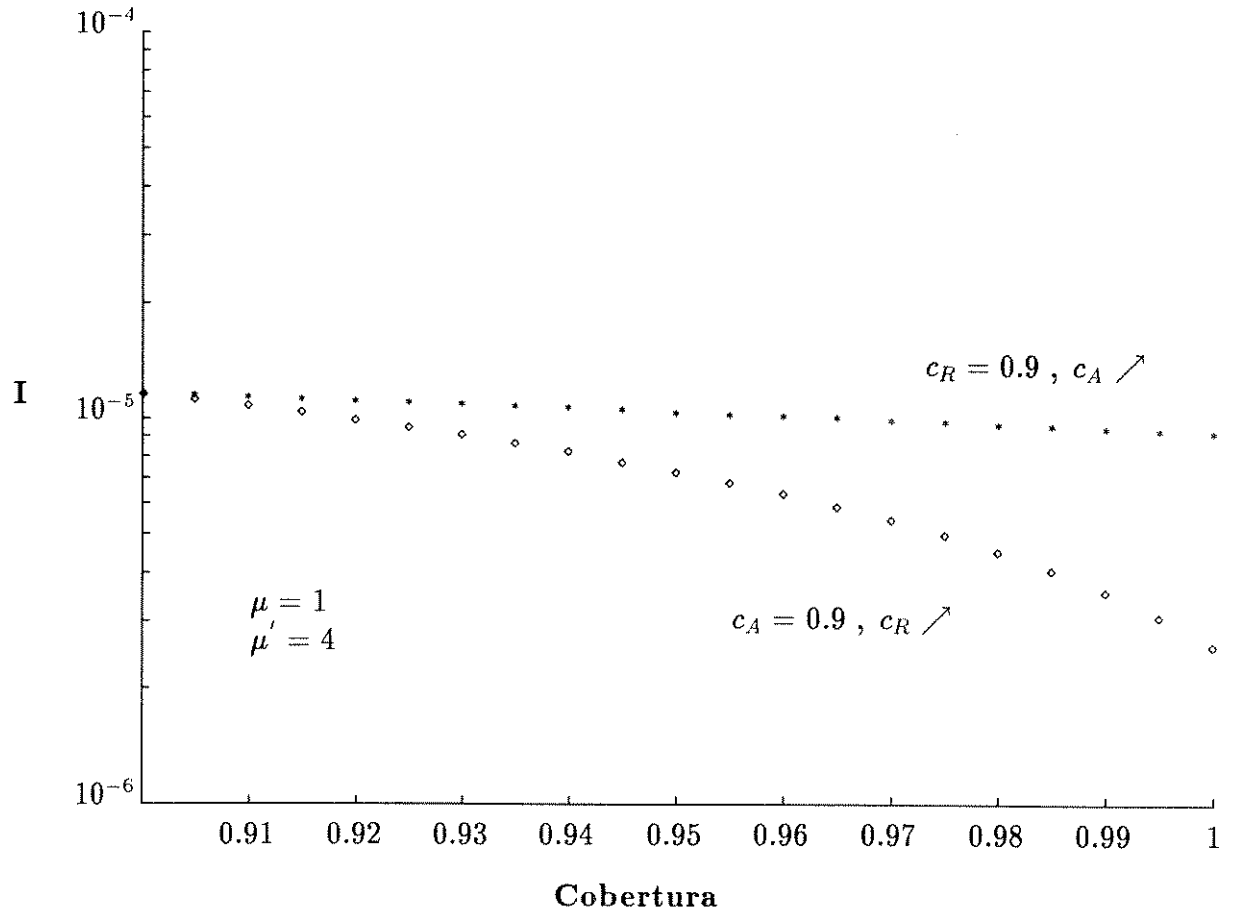


Figura 2.4 - Indisponibilidade de um sistema redundante (*hot standby 1:1*) X cobertura.

A explicação para esta influência maior reside na possibilidade de ocorrência de uma falha não coberta na unidade reserva. Como pode ser observado no modelo da figura 2.3, a ocorrência deste evento leva a uma situação, caracterizada pelo estado 3, a partir da qual só há possibilidade de transição para um estado de falha com taxa de retorno lenta (μ). Por outro lado, apesar da ocorrência de uma falha não coberta da unidade ativa conduzir o sistema a um estado de falha, este estado possui taxa de retorno mais rápida (μ'). Desta forma, em termos comparativos, este último evento tem um impacto menor na indisponibilidade do sistema, o que leva a crer que a existência de uma transição do estado 3 para o estado 2 provocaria uma diminuição desta indisponibilidade.

2.2.4 Influência das supervisões periódicas

A ocorrência de falhas não cobertas nas unidades reservas de sistemas com redundância dinâmica, pode prejudicar sensivelmente o desempenho do sistema em termos de disponibilidade. Isto é explicável a partir do fato de que uma unidade reserva pode ser solicitada a substituir uma unidade ativa em falha, e não se encontrar em condições de realizar tal atividade, posto que está com uma falha não detectada/sinalizada. A incidência com que

este evento pode ocorrer, bem como a magnitude das consequências, reflete os princípios de supervisão adotados na concepção de tais sistemas.

Os princípios de supervisão exercem grande influência na disponibilidade de um sistema, uma vez que determinam o nível de abrangência dos mecanismos de supervisão, a periodicidade das rotinas de testes, as facilidades de sinalização/localização de falhas, a frequência de manutenção, etc. Esta influência tem sido avaliada através de uma gama variada de modelos e abordagens [SIL 90], [YAK 86], [ODA 91] e [HEL 80].

Neste aspecto, a partir do modelo da figura 2.3 pode ser obtida uma avaliação do impacto que a realização de supervisões periódicas na unidade reserva traz à indisponibilidade de um sistema redundante com reconfiguração automática e manual, conforme apresentado em [HOL 91b]. Estas supervisões correspondem a uma sequência de testes realizados periodicamente, e que podem detectar falhas em pontos não cobertos pelos mecanismos de supervisão atuantes ao longo do funcionamento normal do sistema.

Para tanto, é necessário considerar, no modelo da figura 2.3, uma transição do estado 3 para o estado 2 com taxa μ''. A possibilidade de passagem do estado 3 para o estado 2 reflete a utilização das supervisões periódicas na unidade reserva, onde os defeitos/falhas não detectados durante o funcionamento normal desta unidade podem ser percebidos através da sequência de testes que integra o procedimento de supervisão periódica. Com isto, o sistema migra para um estado onde a unidade reserva é caracterizada pela condição de falha detectada, permitindo, por consequência, que o processo de manutenção desta unidade seja realizado.

Do ponto de vista dos procedimentos operacionais, a forma mais adequada de realizar as supervisões periódicas é a intervalos de tempo de regulares. Diante destas circunstâncias, a distribuição do tempo de supervisão, ou seja, do período entre sucessivas supervisões, deve ser considerada como constante. Entretanto, neste caso, a consideração de uma distribuição exponencial não compromete os resultados da avaliação, como pode ser constatado através da análise apresentada no apêndice A. Conforme evidenciam os resultados desta análise, a hipótese de distribuição exponencial conduz a estimativas de indisponibilidade conservativas (se comparadas às obtidas para uma distribuição constante) e é valida, inclusive, para períodos de supervisão da ordem de λ^{-1}.

Assim, é possível descrever o comportamento operacional deste sistema através de um processo markoviano, representado pelo modelo da figura 2.5, o que simplifica enormemente o acesso à expressão e às estimativas de indisponibilidade. Neste modelo, são adotadas as mesmas hipóteses e notação descritas para o modelo da figura 2.3, sendo μ'' a taxa de retorno decorrente da detecção de um falha não coberta, durante a supervisão periódica na unidade reserva. Como a distribuição do tempo de supervisão foi assumida como exponencial, a taxa μ'' corresponde ao inverso do período desta supervisão.

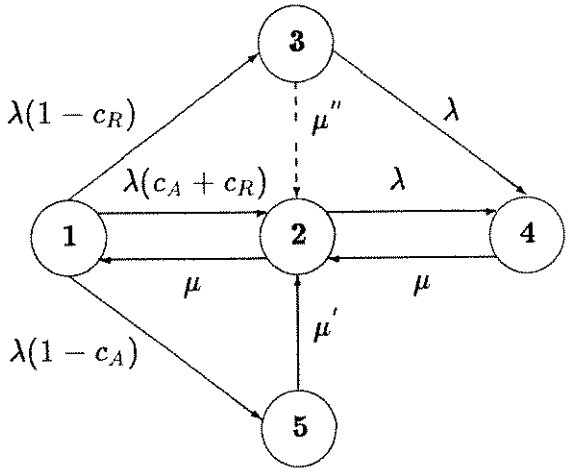


Figura 2.5 - Modelo de Markov para um sistema redundante (*hot standby 1:1*) com coberturas diferenciadas e supervisões periódicas.

Resolvendo o sistema de equações de estados para este modelo, obtem-se a expressão de indisponibilidade I , correspondente à equação (2.41), a qual representa a probabilidade estacionária de ocupação dos estados 4 e 5.

$$\mathrm{I} = \frac{(1-c_A)\lambda/\mu' + 2\lambda^2/\mu^2 + (1-c_R)\lambda^2/\mu(\lambda+\mu'')}{1+2\lambda/\mu + 2\lambda^2/\mu^2 + (1-c_R)\lambda(1+\lambda/\mu)/(\lambda+\mu'') + (1-c_A)\lambda/\mu'} \tag{2.41}$$

Adotando o mesmo procedimento apresentado em [HOL 91b], é possível obter uma análise de sensibilidade do ganho em termos de disponibilidade, com relação aos casos em que $\mu'' = 0$ (I_o) e $\mu'' \neq 0$ (I). Esta relação permite avaliar quanto a indisponibilidade do sistema funcionando sem a supervisão periódica na unidade reserva é superior à indisponibilidade do sistema configurado com esta supervisão.

Considerando que $\frac{2\lambda^2}{\mu^2} << \frac{\lambda}{\mu} << 1, \frac{\lambda}{\mu'} << 1, (1-c_A) \leq 1$ e $(1-c_R) \leq 1$, as expressões de I e I_o podem ser escritas da seguinte forma:

$$\mathrm{I} = \frac{(1-c_A)\lambda/\mu' + (1-c_R)\lambda^2/\mu(\lambda+\mu'')}{1+\lambda(1-c_R)/(\lambda+\mu'')} \tag{2.42}$$

$$\mathrm{I}_o = \frac{(1-c_A)\lambda/\mu' + (1-c_R)\lambda/\mu}{2-c_R} \quad . \tag{2.43}$$

Reescrevendo a equação (2.42), considerando que $\lambda+\mu'' \simeq \mu''$ e $\lambda(1-c_R)+\mu'' \simeq \mu''$, tem-se:

$$\mathrm{I} \simeq \frac{(1-c_A)\lambda/\mu' + (1-c_R)\lambda^2/\mu\mu''}{[\lambda(1-c_R)+\mu'']/\mu''} \simeq (1-c_A)\frac{\lambda}{\mu'} + (1-c_R)\frac{\lambda^2}{\mu\mu''} \quad . \tag{2.44}$$

Assim, a relação I_o/I pode ser determinada dividindo-se (2.43) por (2.44), ou seja,

$$\frac{I_o}{I} = \frac{(1-c_A)\lambda/\mu' + (1-c_R)\lambda/\mu}{(2-c_R)[(1-c_A)\lambda/\mu' + (1-c_R)\lambda^2/\mu\mu'']} =$$

$$= \frac{(1-c_A)/\mu' + (1-c_R)/\mu}{(2-c_R)[(1-c_A)/\mu' + (1-c_R)\lambda/\mu\mu'']} \quad . \tag{2.45}$$

Quando $c_A < 1$ e $\lambda/\mu'' << 1$, o segundo termo do denominador de (2.45) pode ser desprezado, uma vez que é muito menor do que o primeiro. Consequentemente, é possível reescrever I_o/I da seguinte forma:

$$\frac{I_o}{I} \simeq (1 + \frac{1-c_R}{1-c_A}\frac{\mu'}{\mu})/(2-c_R) \quad , \tag{2.46}$$

portanto, neste caso, I_o/I não depende de μ''. Em outras palavras, a existência da supervisão periódica na unidade reserva implica num ganho de disponibilidade, representado pela razão I_o/I, onde, contudo, este ganho não depende (pelo menos em termos significativos) do período da supervisão quando a cobertura na unidade ativa é imperfeita.

De maneira análoga, quando $c_A = 1$, a equação (2.45) pode ser escrita da seguinte forma:

$$\frac{I_o}{I} = \frac{(1-c_R)\mu''}{(2-c_R)\lambda(1-c_R+2\mu''/\mu)} \quad . \tag{2.47}$$

Neste caso, μ'' tem influência direta na relação entre as indisponibilidades I_o e I.

Em síntese, o ganho de disponibilidade proporcionado pelo procedimento de supervisão periódica não depende do período com que a supervisão é realizada, salvo situações onde $c_A \rightarrow 1$ e/ou $\mu'' \rightarrow \lambda$. A tabela 2.2 ilustra estes aspectos através de alguns valores numéricos da relação I_0/I, os quais correspondem aos casos em que $\lambda = 10^{-4}$ *falhas/hora* e $\lambda = 10^{-5}$ *falhas/hora*.

			I_0/I	
c_A	c_R	μ''	$\lambda = 10^{-4}$	$\lambda = 10^{-5}$
0.95	0.95	1/10	4.7	4.8
0.95	0.95	$1/10^2$	4.5	4.7
0.95	0.95	$1/10^3$	3.5	4.6
0.95	0.95	$1/10^4$	1.6	3.5
0.95	0.95	$1/10^5$	-	1.6
1	0.9	1/10	303	3033
1	0.9	$1/10^2$	76.5	758
1	0.9	$1/10^3$	9.9	91

Tabela 2.2 - Ganho de disponibilidade $(I_o/I) \times \mu''$: valores obtidos para $\mu = 1$ e $\mu' = 4$.

As curvas apresentadas na figura 2.6 mostram a influência da cobertura na indisponibilidade do sistema descrito pelo modelo da figura 2.5. Numa das curvas, o c_A é mantido constante, enquanto o c_R assume valores de 0.9 a 1. Na outra curva, a situação se inverte. A confrontação entre estas curvas e as curvas da figura 2.4 permite verificar a diminuição significativa dos valores de indisponibilidade (ganho de disponibilidade), proporcionada pela inclusão do procedimento de supervisão periódica.

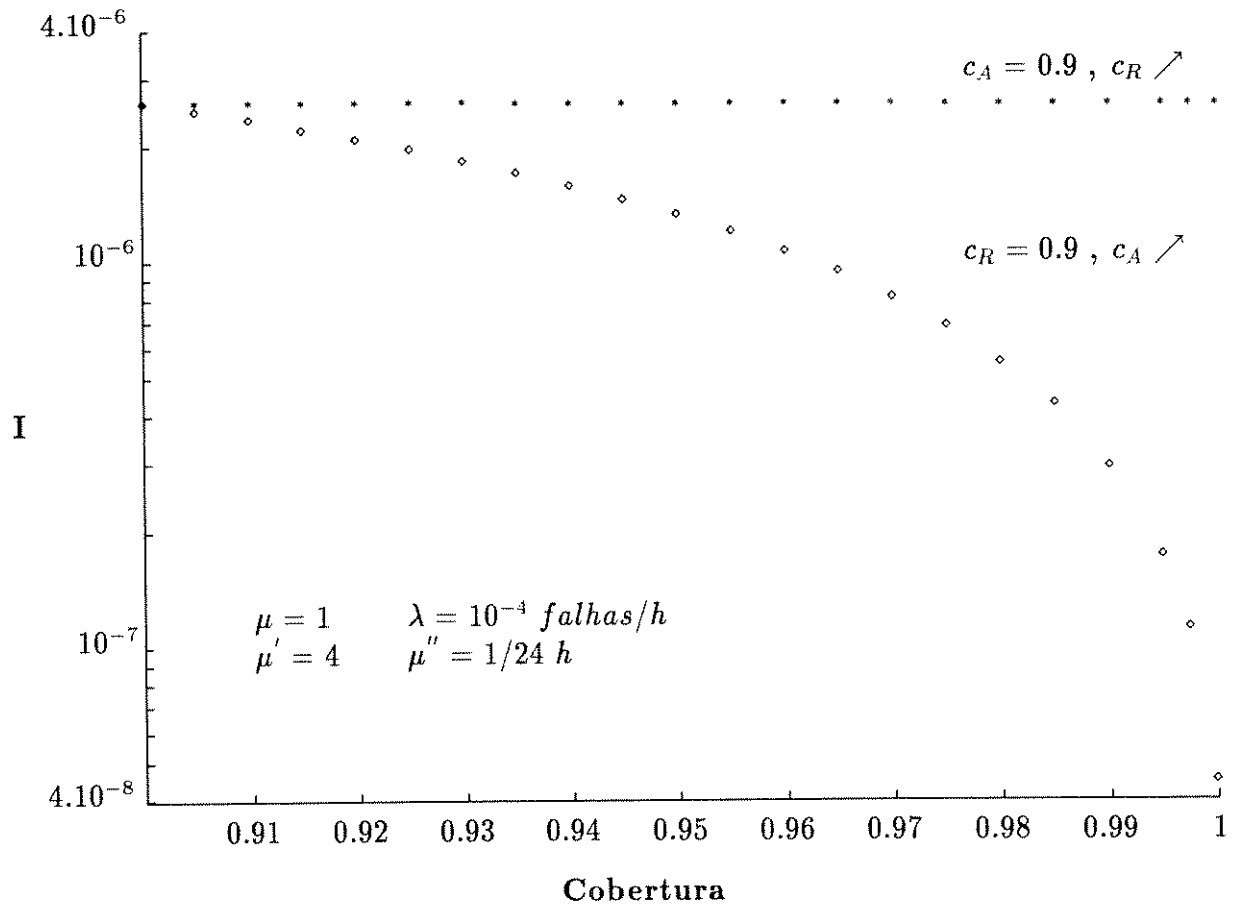


Figura 2.6 - Indisponibilidade de um sistema redundante com supervisões periódicas × cobertura.

Como pôde ser observado até aqui, a formulação matricial e a resolução analítica usual se comportam como procedimentos adequados em termos de sistemas com poucos estados. Para sistemas com elevado número de estados, a solução analítica é praticamente impossível, e mesmo a formulação matricial pode se tornar ineficiente quanto ao esforço computacional exigido. Este aspecto é justificado tanto pela presença de matrizes com grande número de elementos como pelo fato destes elementos apresentarem, frequentemente, ordens de grandeza discrepantes. Visando contornar tais dificuldades, algumas técnicas têm sido empregadas, e.g., as técnicas de agregação e alguns métodos numéricos, as quais são descritas nas subseções a seguir.

2.2.5 Técnica de agregação

A técnica de agregação tem se apresentado como uma opção de particular interesse na solução de modelos de Markov com grande número de estados. A aplicação desta técnica facilita a solução dos modelos, além de atenuar os problemas relacionados com a existência simultânea de taxas de transição com diferentes ordens de grandeza.

A abordagem apresentada por [BOB 86], a qual é resumidamente descrita aqui, contempla modelos de Markov situados neste contexto e, inclusive, proporciona a resolução em regime transiente para modelos com estados absorvedores.

Seja um processo markoviano com um número N de estados, e seja t_m o tempo da missão na qual se deseja determinar as probabilidades de ocupação dos estados. Desta forma é possível agrupar os estados segundo dois subconjuntos:

- Subconjunto $\{E_0\}$ de estados lentos, i.e., estados com todas as transições de saída da ordem de $1/t_m$.
- Subconjunto $\{E_1\}$ de estados rápidos, i.e., estados com pelo menos uma transição de saída cuja taxa é muito maior do que $1/t_m$.

Algoritmo de agregação

O algoritmo proposto consiste na agregação dos estados de $\{E_1\}$, os quais são divididos da seguinte maneira:

- Estados rápidos fechados ou recorrentes $\{C_I\}$, $I = 1, 2, ..., v-1$: desconsiderando as transições lentas, o subconjunto $\{C_I\}$ é dito fechado se todos os seus estados são ergódicos e nenhum estado, não pertencente à $\{C_I\}$, pode ser alcançado a partir de qualquer estado de $\{C_I\}$.
- Estados rápidos transientes $\{C_v\}$: desconsiderando as transições lentas, um estado rápido é chamado transiente se o retorno a este estado é um evento impossível.

Reordenando a matriz Λ de taxas de transição de estados de modo que os estados pertencentes a $\{E_0\}$ sejam numerados de 1 a n_o e os estados pertencentes a $\{E_1\}$ sejam numerados de $n_o + 1$ a $n_o + n_f = N$, e particionando-a em estados lentos e subconjuntos de estados rápidos, a matriz original Λ pode assumir a seguinte forma:

$$\Lambda = \begin{bmatrix} \Lambda_{00} & B_{01} & \dots & B_{0v} \\ B_{10} & D_1 & \dots & B_{1v} \\ \vdots & \vdots & & \vdots \\ B_{v0} & B_{v1} & \dots & D_v \end{bmatrix}$$

onde

Λ_{00}, B_{0I} $(I = 1,2,...,v)$	$\rightarrow$	representam as submatrizes com transições lentas
B_{vI} $(I = 1,2,...,v-1)$	$\rightarrow$	representa as submatrizes que podem conter transições rápidas, se estados rápidos são conectados a $\{C_I\}$ ou a $\{E_0\}$ através de transições rápidas.
D_I $(I = 1,2,...,v)$	$\rightarrow$	representa as submatrizes com pelo menos uma transição rápida em cada linha.
B_{IJ} $(I = 1,2,...,v-1;$ $J = 0,1,...,v, I \neq J)$	$\rightarrow$	representa as submatrizes que contém transições lentas.

Este algoritmo descreve duas abordagens: uma sobre a agregação dos subconjuntos de estados rápidos recorrentes e outra sobre a agregação do subconjunto transiente. A principal diferença entre estes dois tipos de subconjuntos reside no fato de que o tempo de permanência nos estados do subconjunto transiente se aproxima de zero à medida que as taxas de transição rápida tendem a infinito, ao passo que este não é necessariamente o caso do subconjunto recorrente, o qual pode se tornar um conjunto de estados absorvedores. Nestas condições, o tempo de permanência no subconjunto rápido recorrente tende para infinito independentemente dos valores das taxas de transição rápidas e lentas.

Como os sistemas de telecomunicações com aplicações para fins comerciais admitem o processo de reparação e toleram pequenos intervalos de interrupção, os modelos associados a tais sistemas dificilmente incorporam estados classificáveis como rápidos recorrentes. Neste sentido, a seguir é descrita apenas a abordagem envolvendo a agregação do subconjunto de estados transientes. Entretanto, a abordagem para o caso de existência de estados rápidos recorrentes pode ser encontrada em [BOB 86] e [SOU 88].

Agregação dos estados rápidos transientes

Particionando a matriz de transição de estados reordenada, conforme mencionado anteriormente, em quatro submatrizes, pode-se escrever as equações que regem as transições entre estados na seguinte forma:

$$[\dot{\vec{P}}_\sigma \quad \dot{\vec{P}}_v] = [\vec{P}_\sigma \quad \vec{P}_v] \begin{bmatrix} E_{\sigma\sigma} & E_{\sigma v} \\ E_{v\sigma} & D_v \end{bmatrix} \quad , \tag{2.48}$$

onde P_σ representa o vetor das probabilidades de ocupação dos estados lentos e P_v o vetor das probabilidades de ocupação dos estados pertencentes a C_v; $E_{\sigma\sigma}$ compreende as taxas de transição entre os estados lentos; $E_{\sigma v}$ e $E_{v\sigma}$ compreendem as taxas de transição entre os estados lentos e os estados rápidos transientes, e vice-versa; e D_v envolve as taxas de transição entre os estados rápidos transientes.

De (2.48), obtem-se:

$$\frac{d\vec{P}_\sigma(t)}{dt} = \vec{P}_\sigma(t)E_{\sigma\sigma} + \vec{P}_v(t)E_{v\sigma}$$
$$\frac{d\vec{P}_v(t)}{dt} = \vec{P}_\sigma(t)E_{\sigma v} + \vec{P}_v(t)D_v \quad ,$$

Considerando que os estados de C_v já atingiram sua distribuição de estado estacionária, i.e., $\dot{\vec{P}}_v(t) = 0$, é possível obter as seguintes expressões aproximadas:

$$\frac{d\vec{P}^*_\sigma(t)}{dt} = \vec{P}_\sigma E_{\sigma\sigma} - \vec{P}_\sigma E_{\sigma v} D_v^{-1} E_{v\sigma}$$
$$\vec{P}^*_v = -\vec{P}_\sigma E_{\sigma v} D_v^{-1} \quad .$$

Desta forma, a matriz de transição de estados definida sobre o subconjunto de estados lentos é, aproximadamente dada por (2.49).

$$\Lambda^*_\sigma = E_{\sigma\sigma} - E_{\sigma v} D_v^{-1} E_{v\sigma} \quad . \tag{2.49}$$

Interpretanto a equação (2.49), pode-se notar que o primeiro termo do lado direito da igualdade leva em conta as transições entre os estados de $\{E_\sigma\}$, e o segundo termo considera as transições dos estados de $\{E_\sigma\}$ para os estados de $\{C_v\}$, e as transições dos estados de $\{C_v\}$ para os de $\{E_\sigma\}$ regidos pelas probabilidades condicionais do modelo de Markov embutido. Esta probabilidade condicional pode ser melhor explicada através da seguinte relação: uma vez que o processo está num estado de $\{C_v\}$, a probabilidade que ele seja absorvido por um elemento de $\{E_\sigma\}$ é dada por:

$$[D_v]^{-1} E_{v\sigma} \quad .$$

A seguir são apresentados dois exemplos do processo de agregação.

A. Exemplo 1

Consideremos o exemplo apresentado em [BOB 86], o qual compreende um modelo de segurança de funcionamento contendo um submodelo de tratamento de defeitos/erros. Este modelo é representado pelo diagrama de estados da figura 2.7, onde os estados 1 e 2 são estados operacionais com m e $m-1$ itens em funcionamento, respectivamente, e o estado 3 representa um estado de falha do sistema. Diante da ocorrência de um defeito em um determinado item, a qual ocorre com taxa $m\lambda$, o sistema entra no submodelo de tratamento de defeitos/erros.

No estado A deste submodelo, o defeito tanto pode provocar erros e induzir a passagem para o estado E, a uma taxa ρ, como pode ser detectado com taxa δ e o sistema ser reconfigurado para um estado operacional com $m-1$ itens. A taxa de transição de saída

do estado E é dada por k, e u representa a probabilidade do erro ser detectado.

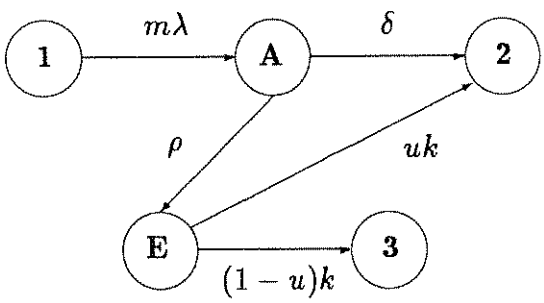


Figura 2.7 - Modelo de Markov incluindo um submodelo de tratamento de defeitos/erros.

Como ρ, δ e k apresentam ordem de grandeza muito superior a λ, os estados A e E formam um subconjunto de estados rápidos transientes, de tal sorte que a matriz Λ pode ser particionada da seguinte forma:

$$\Lambda = \left[\begin{array}{ccc:cc} -m\lambda & 0 & 0 & m\lambda & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hdashline 0 & \delta & 0 & -(\delta+\rho) & \rho \\ 0 & uk & (1-u)k & 0 & -k \end{array}\right]$$

Utilizando a equação (2.49), é possível obter a matriz de taxas de transição A^*_σ do modelo agregado, ou seja,

$$\Lambda^*_\sigma = E_{\sigma\sigma} - E_{\sigma v} D_v^{-1} E_{v\sigma} = \left[\begin{array}{ccc} -m\lambda & m\lambda\frac{\delta+\rho u}{\delta+\rho} & m\lambda(1-u)\frac{\rho}{\delta+\rho} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right] .$$

O modelo de Markov reduzido é mostrado na figura 2.8, onde c é dado por:

$$c = \frac{\delta}{\delta+\rho} + \frac{\rho u}{\delta+\rho} .$$

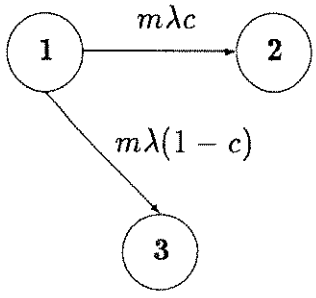


Figura 2.8 - Modelo agregado da figura 2.7.

É interessante notar que c também poderia ser obtido diretamente do modelo da figura 2.7, através de uma simples inspeção, ou seja:

> A transição do estado 1 para o estado A ocorre com taxa $m\lambda$. Considerando a transição de A para 2 governada pelo modelo embutido, i.e., o tempo de permanência nos estados A e E sendo igual a zero, a taxa com que esta transição ocorre é dada pela transição direta de A para 2, com taxa $\delta/(\delta+\rho)$, ou pela transição de E para 2 com taxa $\rho u/(\delta+\rho)$. Desta forma, é possível representar a taxa de transição de 1 para 2 como
>
> $$\lambda_{12} = m\lambda(\frac{\delta}{\delta+\rho} + \frac{\rho u}{\delta+\rho}) = m\lambda c \quad ,$$
>
> onde, na verdade, c é o fator de cobertura do modelo da figura 2.7.

De forma similar, também é possível obter a taxa de transição do estado 1 para o estado 3, a partir da inspeção direta do modelo.

B. Exemplo 2

Consideremos o exemplo de uma estrutura redundante com uma unidade em serviço e duas unidades na reserva, apresentado em [SOU 88], e cujo modelo de Markov é representado pelo diagrama de estados da figura 2.9 . Neste modelo, as unidades possuem taxa de falha λ, c é o fator de cobertura e μ é a taxa de reparo.

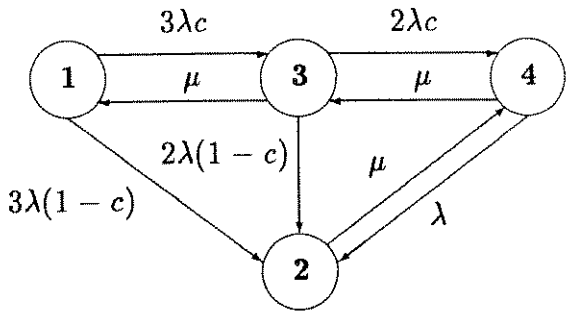


Figura 2.9 - Modelo de Markov para a estrutura redundante 1:2.

Aplicando o procedimento de agregação com o intuito de obter a probabilidade de ocupação do estado 2, tem-se: $\{E_\sigma\} = \{1,2\}$ e $\{C_v\} = \{3,4\}$. Consequentemente, a matriz Λ é particionada da seguinte forma:

$$\Lambda = \left[\begin{array}{cc:cc} -3\lambda & 3\lambda(1-c) & 3\lambda c & 0 \\ 0 & -\mu & 0 & \mu \\ \hdashline \mu & 2\lambda(1-c) & -(2\lambda+\mu) & 2\lambda c \\ 0 & \lambda & \mu & -(\lambda+\mu) \end{array}\right]$$

Utilizando então a equação (2.49), chega-se à seguinte matriz de taxas de transição do modelo agregado:

$$\Lambda_\sigma^* = \left[\begin{array}{cc} -3\lambda + \frac{3\lambda\mu c(\mu+\lambda)}{\Delta} & 3\lambda(1-c) + \frac{3\lambda c[2\lambda(\mu+\lambda)(1-c)+2\lambda^2 c]}{\Delta} \\ \frac{\mu^3}{\Delta} & -\mu + \frac{2\lambda\mu^2(1-c)+\lambda\mu(\mu+2\lambda)}{\Delta} \end{array}\right]$$

onde $\Delta = (\mu + 2\lambda)(\mu + \lambda) - 2\lambda\mu c$.

A partir de Λ_σ^*, as probabilidades de ocupação do estado 2 podem ser facilmente obtidas.

Em síntese, a técnica de agregação é indicada para solucionar dificuldades decorrentes de modelos complexos e da existência de taxas de transição com diferentes ordens de grandeza. Além disso, a agregação propicia, dependendo das proporções do modelo, a obtenção das expressões analíticas dos parâmetros de segurança de funcionamento, conforme mostrado nos exemplos 1 e 2. Entretanto, para modelos com um número de estados extremamente elevado, esta técnica apresenta algumas limitações de ordem computacional, além do que existe a impossibilidade de controle sobre o erro gerado pela aproximação do algoritmo, o que pode torná-la inadequada a avaliações que requeram resultados com alta precisão. Diante deste quadro, alguns métodos numéricos se apresentam como opções capazes de contornar tais limitações.

2.2.6 Técnicas numéricas

A solução numérica para modelos com grande número de estados encontra, normalmente, dois grandes empecilhos computacionais: *i*) os problemas associados à inversão de matrizes de dimensões extremamente altas (algumas vezes maiores do que 10^4) e *ii*) os problemas de estabilidade e eficiência usualmente encontrados na determinação da exponencial de matrizes igualmente grandes, quando se está interessado na análise transiente.

Para o primeiro empecilho, há, como saída, uma gama extensa de pacotes computacionais voltados à solução deste tipo de matriz. Já para o segundo, as alternativas existentes precisam ser bem analisadas, no sentido de observar as vantagens & desvantagens e determinar as que melhor se adaptam às necessidades da avaliação [MOL 78]. Neste particular, Reibman & Trivedi [REI 88] analisam três técnicas numéricas, a saber:

- Uma técnica convencional de solução de equações diferenciais (Runge-Kutta), as quais se apresentam como opções satisfatórias para modelos que não possuem grandes discrepâncias quanto à ordem de grandeza das taxas de transição, e para avaliações com exigências normais de precisão.

- A técnica de randomização, a qual é usualmente mais precisa e eficiente, além de explorar a natureza probabilística do problema e permitir um controle maior do erro.

- Uma técnica que utiliza o método de solução implícita estável (TR-BDF2), a qual é indicada para modelos que possuem taxas de transição com diferentes ordens de grandeza. Entretanto, para que esta técnica apresente resultados com alta precisão existe um alto custo associado.

Em função destas propriedades, a randomização reune algumas vantagens que a torna adequada a muitos tipos de sistema, o que, inclusive, pode ser constatado pelo seu uso crescente nas ferramentas de avaliação de última geração. Baseado nestes aspectos, esta técnica é sumariamente descrita nesta subseção, com o propósito de evidenciar suas principais características.

Técnica de randomização

A técnica de randomização consiste num procedimento numérico de acesso às probabilidades transientes de ocupação dos estados de um processo markoviano, o qual, extensivamente, possibilita também a obtenção de medidas relacionadas a estas probabilidades, e.g., distribuição do tempo de primeira passagem [JEN 53], [GRO 84]. As principais vantagens desta técnica são: a capacidade de manter uma interpretação probabilística do problema, o que não acontece nos métodos numéricos clássicos, e o controle sobre o erro proveniente do tratamento numérico.

Num processo markoviano, a obtenção das probabilidades transientes ocorre com a solução do sistema de equações diferenciais de primeira ordem, representado pela expressão (2.28), ou seja:

$$\dot{\vec{P}}(t) = \vec{P}(t)\Lambda \quad .$$

No sentido de empregar o método de randomização, é adotada a seguinte solução para a equação (2.28):

$$\vec{P}(t) = \vec{P}(0)e^{\Lambda t} \quad , \tag{2.50}$$

onde

$$e^{\Lambda t} = \sum_{n=0}^{\infty} \frac{(\Lambda t)^n}{n!} \quad . \tag{2.51}$$

Normalmente, os procedimentos computacionais para solução de (2.50) conduzem a resultados com erros significativos, como decorrência do fato de Λ possuir elementos diagonais negativos. Sob este aspecto, o método de randomização apresenta outras perspectivas, uma vez que possibilita a utilização de algoritmos que trabalham apenas com números positivos, minimizando, assim, os erros, e permitindo um controle sobre este erro a partir da truncagem da série infinita apresentada em (2.51).

Em linhas gerais, a idéia básica da randomização está em submeter o processo markoviano a um processo de Poisson, segundo um procedimento como o descrito em [ÇIN 75], de modo que a equação (2.50) pode ser solucionada da seguinte forma:

Seja $\{X(t), t \geq 0\}$ um processo de Markov uniformizável, i.e., os elementos diagonais de Λ são uniformemente limitados, de forma que é possível definir a matriz Π como:

$$\Pi = \frac{\Lambda}{\lambda} + I \quad ,$$

onde λ representa um escalar maior ou igual ao maior elemento de Λ (em valor absoluto) e I é a matriz identidade. É possível notar que Π é uma matriz de probabilidades, uma vez que seus elementos, representados por π_{ij}, são tais que $0 \leq \pi_{ij} \leq 1$.

Escrevendo Λ na forma $\Lambda = (\Pi - I)\lambda$, tem-se:

$$e^{\Lambda t} = e^{\Pi\lambda t}.e^{-\lambda t} = \sum_{n=0}^{\infty} \Pi^n \frac{(\lambda t)^n}{n!} e^{-\lambda t} \quad . \tag{2.52}$$

Aplicando (2.52) em (2.50) e truncando a série no termo de ordem N, obtem-se as probabilidades transientes, ou seja,

$$\vec{P}(t) \simeq \vec{P}_o \sum_{n=0}^{N} \Pi^n \frac{(\lambda t)^n}{n!} e^{-\lambda t} \quad . \tag{2.53}$$

O erro devido à truncagem em N pode ser controlado a partir da seguinte relação [GRO 84]:

$$1 - e^{-\lambda t} \sum_{n=0}^{N} \frac{(\lambda t)^n}{n!} \leq \epsilon \quad ,$$

sendo ϵ o fator de controle do erro.

Através da técnica de randomização, também é possível obter expressões computacionais para as probabilidades transientes de múltiplos instantes e para outras medidas transientes, tais como: distribuição do tempo de primeira passagem, tempo de permanência no processo e número esperado de ocorrência de eventos num dado intervalo de tempo. O procedimento para obtenção destas expressões não é tratado aqui, sendo, contudo, encontrado em [GRO 84].

2.3 Panorama Geral das Ferramentas de Avaliação

Esta seção apresenta uma rápida descrição das principais características das ferramentas e técnicas utilizadas recentemente na avaliação da segurança de funcionamento dos sistemas tolerantes a defeitos, além de situá-las dentro do processo evolutivo que tem marcado a história recente destas ferramentas.

Numa descrição deste processo evolutivo, encontrada em [GEI 90], Geist e Trivedi identificam os modelos baseados na análise combinatorial (alguns estendidos de forma a considerar coberturas imperfeitas) como uma primeira geração de ferramentas de avaliação. Um exemplo típico desta geração é o pacote CARE (Computer Aided Reliability Estimation), inclusive a versão modificada CARE II.

A necessidade de avaliações com resultados mais precisos estimulou, rapidamente, o desenvolvimento de ferramentas baseadas em modelos mais elaborados. Assim, surge uma segunda geração de ferramentas calcada em modelos markovianos, a qual permitiu abolir a hipótese de independência, condicionadora dos modelos simplificados da geração anterior. As principais características destas ferramentas, bem como as das gerações subsequentes, são relacionadas a seguir.

2.3.1 Segunda geração

A primeira ferramenta desta geração foi o pacote ARIES (Automated Reliability Estimation System) [WNG 77], o qual é baseado em processos markovianos com estado discreto e tempo contínuo. Nestes processos, a cobertura era considerada instantânea e representada por probabilidades constantes.

O ARIES proporcionou uma grande flexibilidade ao processo de especificação dos modelos e serviu como base para o desenvolvimento futuro de novas ferramentas, apesar de apresentar algumas limitações como a hipótese de taxas de transição constantes.

Uma outra ferramenta desta geração é o SURF (Systeme d' evaluation de la surete de fonctionnement) [LAN 78], a qual foi a primeira ferramenta a considerar a possibilidade de taxas de falha não constantes. Na verdade, isto foi conseguido com a utilização do método dos estágios, que possibilita aproximar uma variável aleatória com distribuição geral, substituindo-a por uma configuração série-paralelo de estágios exponenciais.

Apesar deste método conduzir a modelos com grande número de estados, o SURF mostrou-se eficiente na modelagem de sistemas reparáveis de tamanho moderado. Além disso, constituiu-se numa ferramenta de suma importância, ao viabilizar a representação de taxas de falha dependentes do tempo, bem como taxas de reconfiguração e de reparo.

Uma terceira ferramenta desta geração, também digna de nota, é o CAST (Complementary Analytic Simulative Technique) [GEI 83]. Apesar de possuir um modelo geral que representa um caso especial do ARIES, o pacote CAST contribuiu significativamente no plano conceitual. Este aspecto é justificado pela inclusão do tratamento de defeitos transientes, a qual consistiu na primeira modelagem detalhada da cobertura, embora a hipótese de cobertura instantânea ainda fosse adotada. Além disso, o CAST foi a primeira ferramenta a sugerir a utilização coordenada de simulação nos modelos analíticos.

2.3.2 Terceira geração

A terceira geração de ferramentas ficou caracterizada com o desenvolvimento do CARE III [GEI 83]. O desenvolvimento desta ferramenta foi motivado pelas limitações das abordagens anteriores no que diz respeito à avaliação de sistemas altamente confiáveis, e sedimentou a necessidade de detalhar o processo de tratamento de defeitos/erros na busca por estimativas mais precisas sobre a probabilidade de coberturas imperfeitas. A necessidade de modelos de cobertura detalhados refletia o reconhecimento consensual de que coberturas imperfeitas consistiam em causas primárias de falhas dos sistemas tolerantes a defeitos.

O CARE III considera taxas de transição dependentes do tempo, as quais são classificadas segundo dois grupos: transições rápidas, relacionadas com os mecanismos de tratamento de defeitos/erros, e transições lentas associadas aos mecanismos de ocorrência de defeitos. Os modelos são, então, constituídos por submodelos de ocorrência de defeitos e tratamento de defeitos/erros, onde os primeiros são descritos por modelos de Markov não homogêneos e os segundos por modelos semi-markovianos. Os modelos semi-markovianos são resolvidos separadamente, e os resultados são incorporados ao modelo de ocorrência de defeitos através de métodos numéricos.

A principal limitação desta ferramenta se concentra na impossibilidade de determinar precisamente o erro envolvido nas técnicas de solução dos modelos semi-markovianos. Apesar disto, desempenhou um papel de reconhecida importância, principalmente por fomentar o desenvolvimento de novas ferramentas.

2.3.3 Última geração

Dentre as técnicas e ferramentas correntes, é possível destacar seis pacotes software para avaliação dos parâmetros de segurança de funcionamento, a saber: HARP, SURE, Heiress, SHARPE, Save e SURF-2, os quais são suscintamente descritos a seguir.

A. HARP

Como normalmente acontece, as limitações apresentadas pelas ferramentas em uso corrente, para determinadas aplicações, condicionam a especificação e desenvolvimento de novas técnicas e ferramentas a partir do patamar estabelecido pelas antecessoras. Neste aspecto, algumas aplicações não cobertas de forma satisfatória pelo CARE III, impulsionaram o desenvolvimento do HARP (Hybrid Automated Reliability Predictor) [GEI 90].

O HARP utiliza modelos de Markov não homogenêos na modelagem do processo de ocorrência de defeitos e três técnicas na modelagem da cobertura: modelos semi-markovianos, Redes de Petri Estocásticas Estendidas (as quais permitem o uso de simulação como técnica de resolução), e distribuições empíricas. A modelagem do processo de cobertura é realizada através de dois tipos de submodelos: submodelo de tratamento de defeitos/erros e submodelo de recuperação.

O principal objetivo do HARP é aumentar a flexibilidade na especificação dos processos de tratamento de defeitos/erros, recuperação e ocorrência de defeitos, além de fornecer estimativas conservativas de confiabilidade e reduzir o tempo de execução computacional.

Como limitações, o HARP apresenta problemas de velocidade na resolução de modelos com grande número de estados (da ordem de 25.000 estados). Outro problema envolve a aproximação usada na avaliação de confiabilidade, a qual considera a cobertura como instantânea. Esta aproximação fornece resultados conservativos, conforme desejado, quando são considerados eventos com taxas de transição constantes, porém o mesmo não é verificado para o caso de taxas de transição dependentes do tempo. Este aspecto torna o programa restritivo em termos de algumas aplicações críticas que requerem modelos ainda mais realistas.

B. SURE

O SURE (Semi-Markov Unreliability Range Estimator) [GEI 90] foi desenvolvido, pela NASA, com o objetivo de proporcionar estimativas rápidas e precisas sobre a confiabilidade de sistemas modelados através de processos markovianos com grande número de estados e baixa incidência de falhas.

As estimativas fornecidas pelo SURE apresentam caráter conservativo, são obtidas de forma mais rápida do que as do HARP, porém com menor precisão. Mesmo assim, a precisão do SURE é suficiente para as necessidades de projeto, e atendem aos objetivos para os quais foi idealizado.

Como limitação desta ferramenta, destaca-se a impossibilidade de inclusão da hipótese de dependência de tempo global. Consequentemente, esta restrição pode conduzir a resultados imprecisos, em termos de modelagem de alguns sistemas que requeiram a hipótese de taxas de falha variantes com o tempo, e.g, aqueles com aplicações em controle de vôo e embarcações tripuladas. Para estes sistemas, fatores como variações ambientais durante o tempo de missão e o envelhecimento dos componentes constituem aspectos relevantes,

de maneira que a não consideração dos mesmos, como fatores determinantes de taxas de falha dependentes do tempo, pode não levar a estimativas conservativas, as quais são indispensáveis à avaliação da segurança de funcionamento de tais sistemas.

C. Heiress

O software Heiress (Hierarchical estimation of internal reliability by skewed sampling) foi especificado com o propósito de fornecer estimativas de confiabilidade altamente precisas, a partir de modelos markovianos e semi-markovianos. Este software utiliza a técnica de redução de variância, denominada *importance sampling* [GEI 90], a qual permite o uso de simulação, inclusive quando se requer resultados com alta precisão.

Entretanto, novas versões desta técnica tornam-se necessárias no sentido de melhorar o desempenho da simulação. Além disso, o emprego de simulação nos modelos completos de confiabilidade (i.e., modelo de ocorrência de falhas, incluindo os submodelos de tratamento de defeitos/erros) ainda se apresenta como uma alternativa bastante custosa, sendo, todavia, mais indicado para os submodelos de tratamento de defeitos/erros.

D. SHARPE

A proposta básica do SHARPE (Symbolic Hierarchical Automated Reliability and Performance Evaluator) é fornecer um *kit* de ferramentas que inclui diferentes tipos de modelo. Esta diversificação de modelos abrange desde modelos comuns de segurança de funcionamento, e.g., diagramas em blocos e árvores de falha [SAH 87], até modelos markovianos e semi-markovianos, além de modelos de acesso a parâmetros de desempenho como, por exemplo, redes de filas, e modelos que permitem a avaliação combinada de parâmetros de desempenho e confiabilidade.

A principal vantagem desta ferramenta é que o usuário pode escolher o tipo de modelo (bem como combinações de modelos) que melhor se adapta à sua aplicação.

E. Save

O Save (System Availability Estimator) [GOY 86] proporciona estimativas da disponibilidade (estacionária e transiente) de sistemas modelados através de processos markovianos com tempo contínuo. A principal vantagem deste pacote está na versatilidade da linguagem de formulação dos dados de entrada, a qual , combinada com a técnica de armazenagem esparsa aplicada sobre a matriz de taxas de transição, e a subsequente resolução do sistema de equações de transição de estados pela técnica de randomização, torna esta ferramenta eficiente para a avaliação de disponibilidade de sistemas com elevado número de estados.

F. SURF-2

Através do software SURF-2 [ARL 89a], o usuário tem acesso a algumas medidas de segurança de funcionamento, a partir de dois tipos básicos de modelos: modelos de Markov e Redes de Petri Estocásticas. As medidas fornecidas por este software compreendem: pro-

babilidades de ocupação dos estados, tempo médio de permanência nos estados e algumas medidas relativas a custos.

Através das probabilidades de ocupação dos estados, os usuários podem ter acesso à confiabilidade, disponibilidade, segurança contra violações (*security*) e mantenabilidade, todas em regime transiente. Além disso, também é fornecida a disponibilidade em regime estacionário.

Dentre os tempos de permanência, o SURF-2 fornece as seguintes medidas: MTFF, MTTF, MTBF, MDT e MUT. No que tange às medidas de custos, estas estão associadas a taxas de rendimento e bonificações referentes à ocupação dos estados dos modelos e suas respectivas transições.

Quanto aos métodos de resolução, o SURF-2 utiliza duas possibilidades em função do tipo de medida: para sistemas de equações lineares, destinados à obtenção das probabilidades em regime estacionário e dos tempos médios de permanência nos estados, são utilizados métodos iterativos de resolução; ao passo que para sistemas de equações diferenciais de primeira ordem, voltados para a determinação das probabilidades transientes, é utilizada a técnica de randomização.

2.4 Conclusões

Este capítulo reune as técnicas e ferramentas mais usadas na avaliação dos parâmetros de segurança de funcionamento, com enfoque voltado para os defeitos/falhas de origem hardware. Através das técnicas e ferramentas apresentadas, é possível avaliar uma sucessão muito ampla de sistemas computacionais, entre eles, os sistemas comerciais de telecomunicações, contemplando-se as seguintes necessidades:

- Acesso aos parâmetros de segurança de funcionamento;
- Conhecimento detalhado sobre determinadas características operacionais dos equipamentos e sistemas;
- Detecção de particularidades dos sistemas, cujo impacto nos parâmetros de segurança de funcionamento é significativo; etc.

Além disso, a diversificação de ferramentas implica na existência de opções com diferentes níveis de complexidade, as quais podem ser escolhidas de forma a conciliar as exigências da avaliação, ou seja, precisão dos resultados, tempo de resposta, custos, etc. Estes aspectos tornam-se mais evidentes com as aplicações proporcionadas pelos estudos de caso mostrados nos capítulos posteriores.

Apesar de apresentarem um caráter multipropósito, estas ferramentas, enquanto facilidades de prateleira, não contemplam todas as situações possíveis, de forma que é comum a necessidade de métodos e ferramentas específicas, moldados por condições de contorno impostas por políticas de desenvolvimento e características particulares de determinados

sistemas. Neste sentido, o domínio de técnicas de avaliação da segurança de funcionamento possibilita a elaboração de ferramentas voltadas às necessidades de projeto, ou mesmo a adaptação das já existentes.

A avaliação de indisponibilidade realizada no SAMSAT, a qual é reportada no capítulo 3, inclui um exemplo elucidativo destas necessidades. Nesta avaliação, conforme pode ser observado, foi desenvolvida uma metodologia para levantamento do fator de cobertura, com o intuito de proporcionar uma avaliação mais realista e, ao mesmo tempo, atender a necessidades específicas do projeto como: utilização dos recursos disponíveis (evitando o desenvolvimento extra de subprodutos) e rapidez na obtenção do levantamento.

Em termos de aplicações práticas, um outro aspecto importante da avaliação de segurança de funcionamento, como suporte técnico-gerencial ao desenvolvimento de sistemas, pode ser observado através da avaliação do PSN COMPAC, apresentada no capítulo 4. A partir dos resultados fornecidos por esta avaliação, foi possível sugerir alternativas que contribuem para elevar o nível de segurança de funcionamento do sistema e suas partes, como, por exemplo, a adoção de um simples procedimento operacional programado.

Capítulo 3

Avaliação do SAMSAT

O SAMSAT é um sistema de comunicação de dados/voz via satélite utilizando a técnica AMDT (Acesso Múltiplo por Divisão de Tempo) com multifrequência [GON 86]. Este sistema fornece capacidade de comunicação a baixas e médias taxas [PIT 89], e foi desenvolvido pelo CPqD-TELEBRÁS com participação de indústrias do parque nacional.

No que diz respeito a um sistema de comunicação de dados via satélite, a qualidade de serviço percebida pelos usuários está fortemente vinculada à capacidade do sistema em fornecer facilidades de comunicação, conforme as necessidades de seus usuários. Neste sentido, a disponibilidade dos equipamentos e de um enlace entre dois usuários consiste no atributo de segurança de funcionamento que melhor traduz as impressões do usuário sobre o funcionamento do sistema.

Em termos de requisito de segurança de funcionamento para este tipo de sistema, o CCIR, através da recomendação 579-1 [CCI 86], considera que a disponibilidade do caminho digital referente ao estabelecimento de um enlace é determinado pelos efeitos combinados da disponibilidade dos equipamentos e do meio de propagação. Em tal contexto, o CCIR recomenda que a indisponibilidade de um enlace, devida aos equipamentos envolvidos, não seja maior do que 0.2% de um ano.

Além disso, num sistema como o SAMSAT, as funções de controle e supervisão, vitais ao funcionamento de todo o sistema, são realizadas de modo centralizado. Neste caso, a indisponibilidade do hardware responsável por estas funções representa um outro indicativo da segurança de funcionamento, igualmente importante. Desta forma, a avaliação aqui apresentada é direcionada no sentido de fornecer estimativas de indisponibilidade segundo dois níveis:

- Indisponibilidade sistêmica, a qual abrange:
  - Indisponibilidade da estrutura responsável pelas funções de controle e supervisão do sistema (ERC),
  - Indisponibilidade dos equipamentos que compõem esta estrutura;
- Indisponibilidade percebida pelo usuário, compreendendo:
  - Indisponibilidade dos equipamentos aos quais os usuários estão conectados,

– Indisponibilidade dos equipamentos envolvidos num enlace entre dois usuários.

O enfoque adotado nesta análise visa contemplar a indisponibilidade devida a falhas de origem hardware e, sobretudo, realçar os aspectos relacionados com a cobertura. No SAMSAT, a cobertura de defeitos assume um papel relevante, uma vez que o sistema apresenta controle centralizado, conforme descrito posteriormente. Este aspecto leva à necessidade de avaliações mais abrangentes em termos da cobertura. Neste particular, a análise da cobertura contribui tanto para a obtenção de estimativas mais precisas sobre os parâmetros de segurança de funcionamento, como para o fornecimento de informações sobre a atuação dos mecanismos de supervisão e, consequentemente, sobre a utilização das partes redundantes.

O objetivo deste capítulo é apresentar as fases do processo de avaliação da segurança de funcionamento do SAMSAT, e propor uma metodologia para avaliação da cobertura. Através deste processo, é possível acompanhar todo o ciclo de avaliação de um sistema em desenvolvimento, e, ao mesmo tempo, ilustrar a aplicação de algumas técnicas descritas no capítulo 2. A metodologia para levantamento do fator de cobertura foi elaborada com o intuito de proporcionar uma maior precisão nas estimativas dos parâmetros de segurança de funcionamento, sem, contudo, elevar os custos do projeto. Para tanto, este capítulo está estruturado da seguinte forma:

A seção 3.1 apresenta uma rápida descrição sobre as características sistêmicas do SAMSAT, bem como sua arquitetura hardware, princípios de funcionamento e filosofia de tolerância a defeitos. A seção 3.2 descreve o processo de modelagem que permite avaliar a indisponibilidade da estrutura hardware do SAMSAT, evidenciando o impacto da cobertura na indisponibilidade das partes redundantes do sistema. A seção 3.3 mostra a metodologia adotada no levantamento do fator de cobertura dos equipamentos redundantes do SAMSAT. A seção 3.4 apresenta estimativas da indisponibilidade para os parâmetros avaliados, realçando os resultados mais significativos. A seção 3.5 destaca alguns comentários conclusivos sobre a avaliação como um todo.

3.1 Descrição do SAMSAT

O SAMSAT é constituído por Estações Terminais Satélite (ETS), as quais fornecem portas de acesso aos serviços oferecidos pelo sistema [GON 86], e por duas Estações de Referência: uma principal (ER0) e outra secundária (ER1), conforme mostra a figura 3.1. As Estações de Referência localizam-se em pontos geográficos distintos e são instaladas no domínio da empresa operadora, o que implica em assistência local quanto à operação e manutenção.

A ER0 e a ER1 são responsáveis pela realização centralizada das funções de controle e supervisão que coordenam a aplicação da técnica AMDT. Estas funções compreendem o fornecimento da referência de tempo para o sistema, e a execução de procedimentos de controle que objetivam garantir o sincronismo das estações, ou seja, garantir que os surtos de dados transmitidos pelas estações não se sobreponham ao formarem o quadro AMDT [GON 86], como pode ser visualizado através da figura 3.1. Em síntese, o desempenho

correto destas funções é essencial à integridade dos enlaces estabelecidos, sendo, portanto, vital ao funcionamento do sistema.

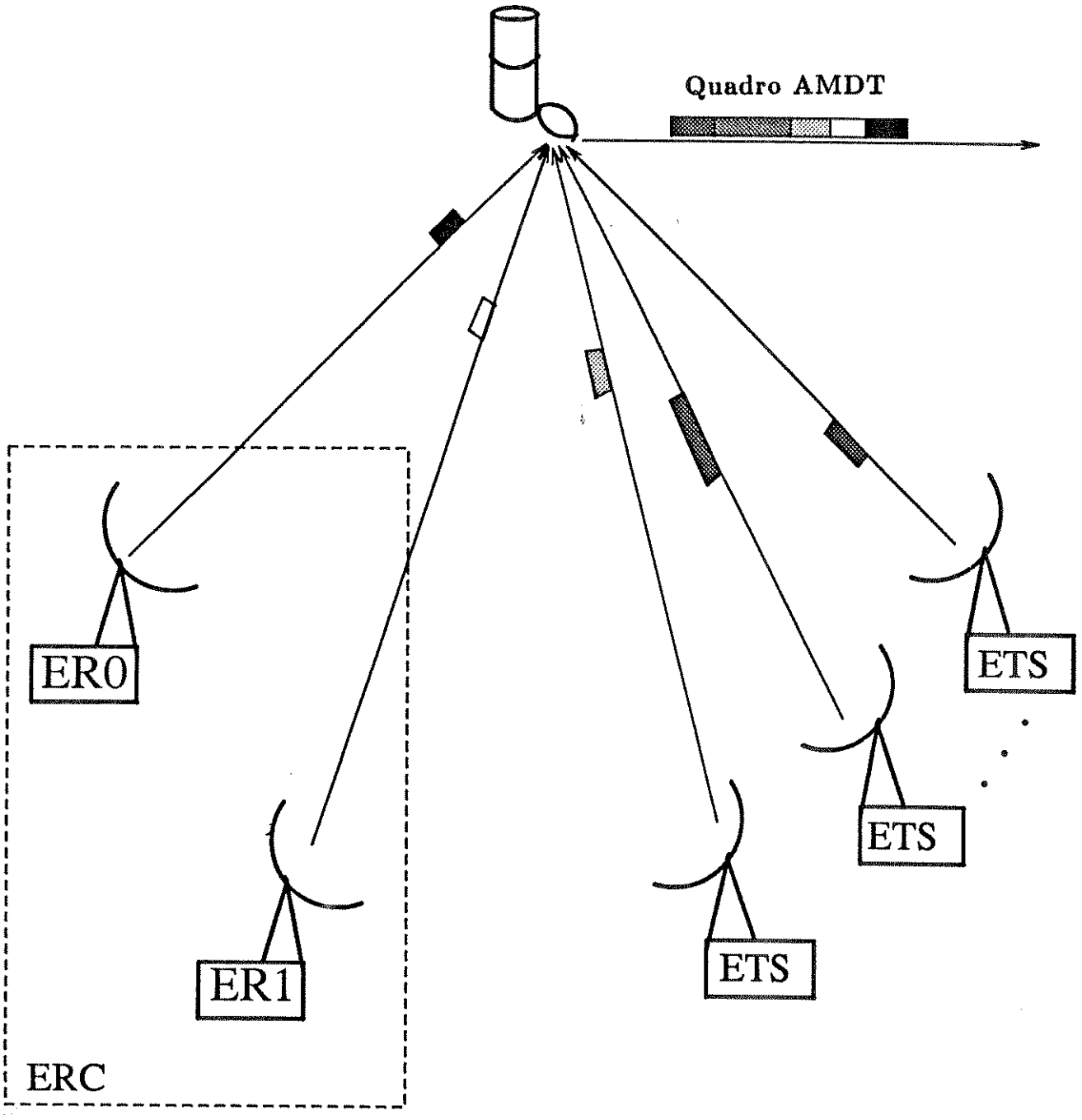


Figura 3.1 - Arquitetura e Estrutura de Sincronismo do SAMSAT.

As Estações de Referência operam de forma conjunta, formando uma estrutura redundante de referência e controle (ERC) que obedece ao princípio de redundância dinâmica *hot standby*. Esta estrutura redundante tem como objetivo atribuir um elevado nível de

segurança de funcionamento ao desempenho das atividades essenciais à manutenção do sincronismo do sistema [FRA 89]. A ER0 é preferencialmente a referência ativa, enquanto a ER1 é mantida na reserva, em condições de substituir a ER0 no caso dela sair de operação. Esta mudança de referência ativa ocorre sem prejuízos para os enlaces estabelecidos.

Ainda em consonância com o propósito de garantir a qualidade de serviço oferecida pelo sistema, cada Estação de Referência e as ETS's configuradas na versão redundante (ETS's redundantes) são dotadas de redundância nas partes essenciais ao desempenho de suas funções, ou seja, nos Equipamentos de RF (ERF), e no Equipamento de Banda-Básica (EBR e EBT redundante, pertencentes, respectivamente, à ER0/ER1 e à ETS redundante).

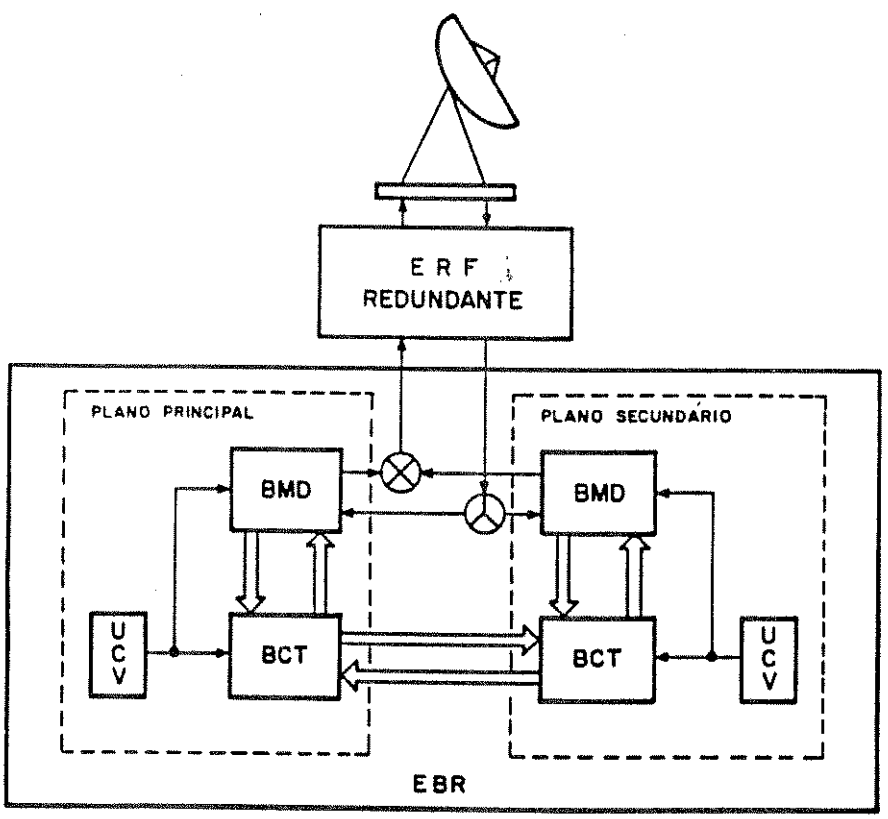


Figura 3.2 - Diagrama em blocos da ER0/ER1.

Como pode ser visto nas figuras 3.2 e 3.3, tanto o EBR como o EBT redundante possuem dois planos de controle (denotados simplesmente por planos): um principal e outro secundário. Cada plano é composto por um Bloco de Controle (BCT), um Bloco Modem (BMD) e uma unidade responsável pela alimentação (UCV). Os planos podem passar de ativo a reserva e vice-versa, e o plano que se encontra como reserva é atualizado frequentemente pelo ativo. Estas atualizações deixam o plano reserva informado sobre as condições operacionais do sistema, de forma que uma mudança de plano transcorre sem interrupções de funcionamento do EBR/EBT redundante.

A ETS não redundante proporciona uma configuração mais econômica para o usuário, uma vez que sua estrutura hardware é simplificada pelo fato dos seus equipamentos (ERF e

EBT) não apresentarem redundância, como pode ser observado na figura 3.4. Mesmo nesta configuração, o EBT possui duas UCV operando redundantemente, no sentido de garantir uma alimentação mais confiável. Ambos os EBT's, o redundante e o não redundante, possuem a mesma configuração quanto às unidades de linha (ULV/ULG), as quais são responsáveis pelas interfaces físicas e elétricas do usuário.

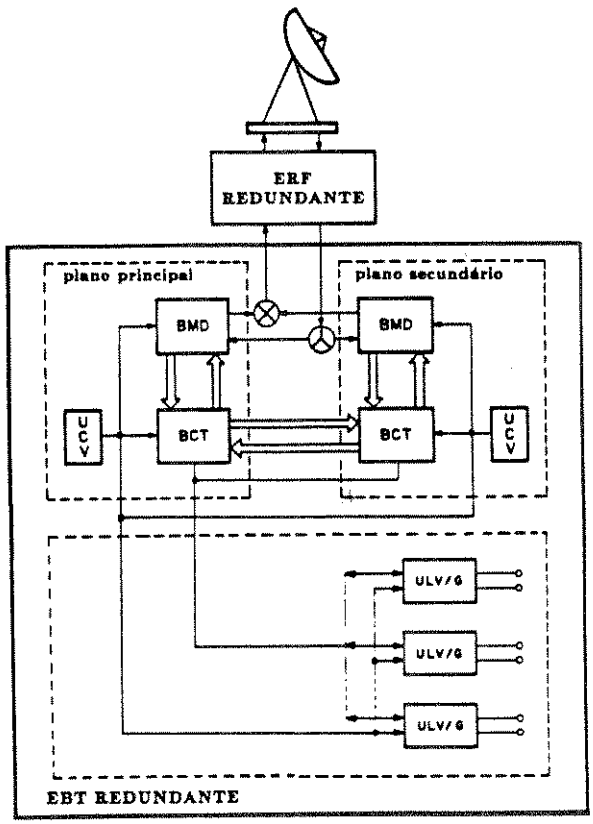


Figura 3.3 - Diagrama em blocos da ETS redundante.

A composição e aspectos de funcionamento dos ERF não são abordados aqui por não se enquadrarem no escopo da análise. Entretanto, este assunto pode ser encontrado em [GON 86] e [GON 87], e, adicionalmente, dados de disponibilidade referentes a equipamentos de RF utilizados em estações terrestres de sistemas comerciais de comunicação por satélite são encontrados em [FEI 86].

Mediante a ocorrência de falhas, as atividades de mudança de plano e/ou referência ativa são viabilizadas por mecanismos de supervisão, destinados à detecção/localização de defeitos e à realização dos procedimentos de reconfiguração entre as partes redundantes. Estes mecanismos são constituídos por dispositivos hardware/software, e fazem parte da filosofia de supervisão empregada no SAMSAT [GON 87].

Conforme abordado nos capítulos 1 e 2, a eficiência destes mecanismos de supervisão, frente ao surgimento de falhas, é indicada através do fator de cobertura. No caso da arquitetura redundante empregada no SAMSAT, esta eficiência é representada por dois fatores de cobertura: c, no que concerne aos mecanismos responsáveis pela mudança de plano do EBR/EBT redundante, e c', referente aos mecanismos responsáveis pela mudança de

referência ativa. Os aspectos de modelagem relativos a c e c' são apresentados na próxima seção.

Os mecanismos de supervisão atuam no plano independentemente destes se encontrarem na condição ativa ou reserva, possuindo praticamente o mesmo nível de abrangência em ambas condições. É evidente que, durante a permanência do plano na condição reserva, o desempenho dos mecanismos de supervisão compreende apenas as atividades de detecção e localização de defeitos, uma vez que não há necessidade de reconfiguração, dado que o plano não se encontra na condição ativa.

Além destas funções, os mecanismos de supervisão fornecem informações, obtidas por sinalização local e remota, que auxiliam no processo de diagnose de falhas/defeitos. Através destas informações é possível reportar as condições operacionais de todas as estações do sistema, proporcionando subsídios valiosos à implantação de estratégias de reparo eficientes.

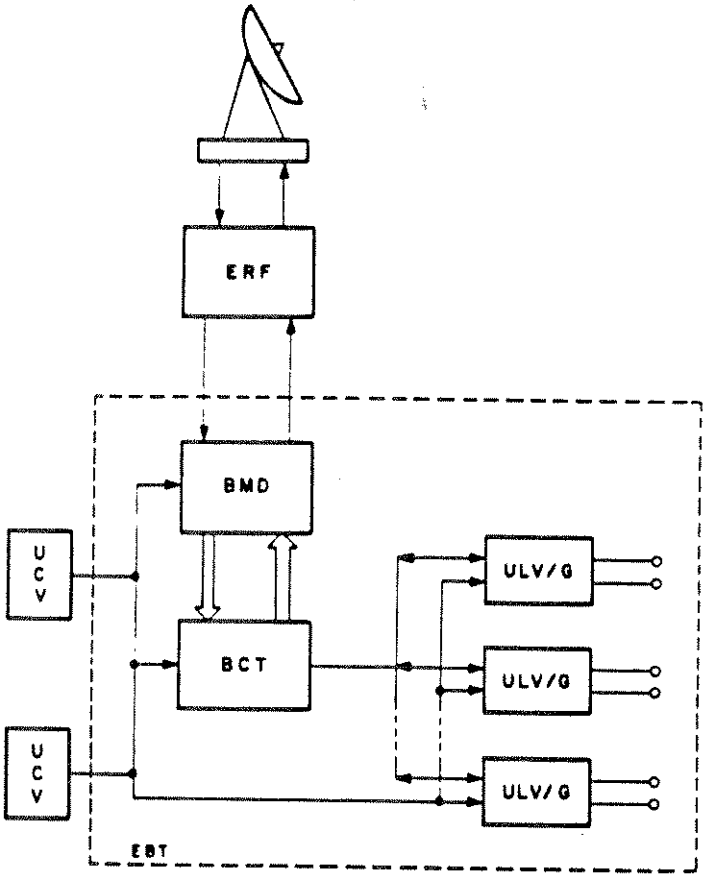


Figura 3.4 - Diagrama em blocos da ETS.

Como pode ser notado, o desempenho integrado de todas as funções dos mecanismos de supervisão está diretamente relacionado com a utilização eficiente das arquiteturas redundantes das estações e, por extensão, com a manutenção do sincronismo do sistema. Estes aspectos evidenciam a importância da atuação adequada dos mecanismos de supervisão (a

qual se reflete no fator de cobertura), no que se refere à obtenção de uma alta disponibilidade dos equipamentos e estrutura de controle. Consequentemente, a análise da cobertura é um dos pontos primordiais na avaliação de indisponibilidade deste sistema.

3.2 Modelos de Segurança de Funcionamento

A modelagem do comportamento operacional dos equipamentos redundantes e da estrutura de controle, em termos do processo de falha e reparação, proporciona a obtenção das seguintes informações relacionadas a estes equipamentos e estrutura:

- Estimativas de indisponibilidade;
- Efeitos da cobertura nos parâmetros de segurança de funcionamento, mais especificamente na indisponibilidade;
- Indicação sobre a necessidade de ampliar ou reprojetar as facilidades de detecção & localização e de reconfiguração, responsáveis pela supervisão de pontos críticos que possam prejudicar a segurança de funcionamento do sistema e suas partes.

Os modelos de acesso às estimativas de indisponibilidade, a partir das quais é possível extrair as informações mencionadas, são apresentados a seguir. Estes modelos descrevem o processo de falha e reparação do EBR/EBT redundante e da ERC através de um processo markoviano com estado discreto e tempo contínuo.

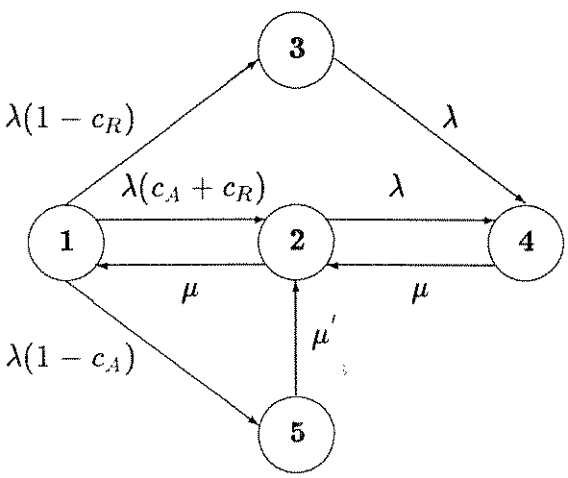
3.2.1 Modelo para o EBR/EBT redundante

Estimativas de indisponibilidade hardware do EBR/EBT redundante (parte de modem e controle) podem ser obtidas a partir do modelo da figura 3.5, o qual é encontrado em [HOL 91a]. Este modelo apresenta uma descrição simplificada destes equipamentos, visando a tratabilidade analítica da solução, a qual é almejada com o propósito de facilitar a observação dos efeitos da cobertura na indisponibilidade dos referidos equipamentos. É importante ressaltar que a simplificação adotada para este modelo apresenta um caráter conservativo, o que resguarda os resultados obtidos.

No modelo da figura 3.5, a cobertura é considerada instantânea, relacionada com falhas hardware permanentes, e compreende a probabilidade de sucesso associada aos processos de detecção (c_d), localização (c_l) e reconfiguração (c_r), i.e., $c = c_d c_l c_r$. A cobertura é diferenciada em função do plano se encontrar na condição ativa ou reserva. No caso do plano reserva, a cobertura é dada pela probabilidade de sucesso das atividades de detecção e localização, uma vez que não há necessidade de reconfiguração. Na verdade, este aspecto permite assegurar que, se os mecanismos de supervisão atuam de forma idêntica independentemente do plano ser ativo ou reserva, a cobertura do reserva é maior ou igual ao do ativo, uma vez que a probabilidade condicional de sucesso da cobertura do reserva não inclui a probabilidade de sucesso de reconfiguração (c_r). As demais hipóteses relacionadas a este modelo são:

i) É assumido que a ocorrência de uma falha não coberta do plano ativo gera indicações de alarme suficientes à caracterização da condição de falha, de forma que é possível a reconfiguração manual entre o plano ativo e o reserva, a uma taxa μ'.

ii) É considerado que, após um reparo, o sistema volta a uma condição operacional e que há apenas um reparador com taxas μ e μ'.



Estado 1 → Dois planos do EBR/EBT redundante em funcionamento.

Estado 2 → Um plano do EBR/EBT redundante em falha, mas o equipamento continua em funcionamento.

Estado 3 → Falha não coberta do plano reserva do EBR/EBT redundante. O equipamento continua em funcionamento.

Estado 4 → Falha de ambos os planos, consequentemente, falha do equipamento.

Estado 5 → Falha não coberta do plano ativo, o que caracteriza uma falha do equipamento.

Figura 3.5 - Modelo de Markov para o EBR/EBT redundante (parte de modem e controle).

Este modelo obedece à seguinte notação:

λ → taxa de falha de cada plano,

c_A → fator de cobertura relacionado com o plano ativo do EBR/EBT redundante,

c_R → fator de cobertura relacionado com o plano reserva do EBR/EBT redundante,

$\mu \rightarrow$ taxa de reparo de um plano (falhas permanentes), $\mu = 1/MTTR$,

$\mu' \rightarrow$ taxa de reparo decorrente de uma reconfiguração manual. $\mu' > \mu$.

Com relação ao procedimento de reconfiguração manual, quando aplicado ao EBT redundante, é importante ressaltar que este equipamento pode ser instalado nas dependências do usuário, e, em tais circunstâncias, não é assistido localmente em termos de operação & manutenção. Desta forma, para que o modelo da figura 3.5 seja igualmente válido para o EBR e EBT redundante, é necessário que o usuário deste equipamento se encontre capacitado a realizar o procedimento de reconfiguração manual, no sentido de assegurar que μ' seja maior do que μ.

A título de ilustração do emprego das técnicas de avaliação apresentadas no capítulo 2, a obtenção das probabilidades de ocupação dos estados do modelo é feita de duas maneiras: uma, analítica, e outra, numérica. A analítica visa destacar qualitativamente os efeitos da cobertura na indisponibilidade dos equipamentos. A numérica apresenta uma avaliação quantitativa destes efeitos.

A partir da solução do sistema de equações de probabilidade estacionária de ocupação dos estados, considerando $c_A = c_R = c$, é possível obter a expressão da indisponibilidade (I), contabilizando-se as probabilidades P_4 e P_5, relativas aos estados 4 e 5, as quais são dadas por:

$$P_4 = \frac{2\lambda^2/\mu^2 + (\lambda/\mu)(1-c)}{2 - c + (\lambda/\mu)(3-c) + 2\lambda^2/\mu^2 + (\lambda/\mu')(1-c)} \tag{3.1}$$

$$P_5 = \frac{(\lambda/\mu')(1-c)}{2 - c + (\lambda/\mu)(3-c) + 2\lambda^2/\mu^2 + (\lambda/\mu')(1-c)} \quad . \tag{3.2}$$

Consequentemente,

$$\mathrm{I} = \frac{2\lambda^2/\mu^2 + (\lambda/\mu)(1-c) + (\lambda/\mu')(1-c)}{2 - c + (\lambda/\mu)(3-c) + 2\lambda^2/\mu^2 + (\lambda/\mu')(1-c)} \quad . \tag{3.3}$$

Como $\lambda/\mu' << 1$, $\lambda/\mu << 1$ e $\lambda^2/\mu^2 << \lambda/\mu$, é possível simplificar a equação (3.3) com o intuito de tornar perceptível a influência de c sobre I. Para tanto, analisemos os casos em que $c = 1$ e $c = 0.5$. Para $c = 1$, a indisponibilidade dada pela expressão (3.3), denotada por $\mathrm{I}_{c=1}$, pode ser aproximada para $\mathrm{I}_{c=1} \simeq 2\lambda^2/\mu^2$, ao passo que para $c = 0.5$, a indisponibilidade ($\mathrm{I}_{c=0.5}$) pode ser simplificada para:

$$\mathrm{I}_{c=0.5} \simeq \frac{\lambda}{3\mu}(1 + \frac{\mu}{\mu'}) \quad .$$

Considerando-se $\mu/\mu' = 1/2$, tem-se que $\mathrm{I}_{c=0.5} = \lambda/2\mu$. Comparando, então, os dois casos, temos:

$$\frac{\mathrm{I}_{c=0.5}}{\mathrm{I}_{c=1}} = \frac{\lambda/2\mu}{2\lambda^2/\mu^2} = \frac{\mu}{4\lambda} \quad .$$

Esta diferença pode ser melhor percebida quando assumimos valores típicos para λ e μ. Considerando $\lambda = 1.72 \times 10^{-4} falhas/h$ e $\mu = 1$ $(MTTR = 1h)$, podemos notar que,

$$\frac{I_{c=0.5}}{I_{c=1}} \simeq 1453 \ ,$$

ou seja, a indisponibilidade para o caso em que $c = 0.5$ é aproximadamente 1453 vezes maior do que a do caso em que $c = 1$.

As curvas da figura 3.6, a qual foi extraída de [HOL 91a], mostram mais detalhadamente o impacto da cobertura na indisponibilidade do EBR/EBT redundante (parte de modem e controle). Estas curvas mostram a indisponibilidade em função do MTTR, correspondendo a três valores de c, e foram obtidas considerando-se $\mu' = 2$ e $\lambda = 1.72 \times 10^{-4} falhas/h$. Adicionalmente, é possível notar o impacto da estratégia de reparo, refletida através do MTTR, na indisponibilidade destes equipamentos.

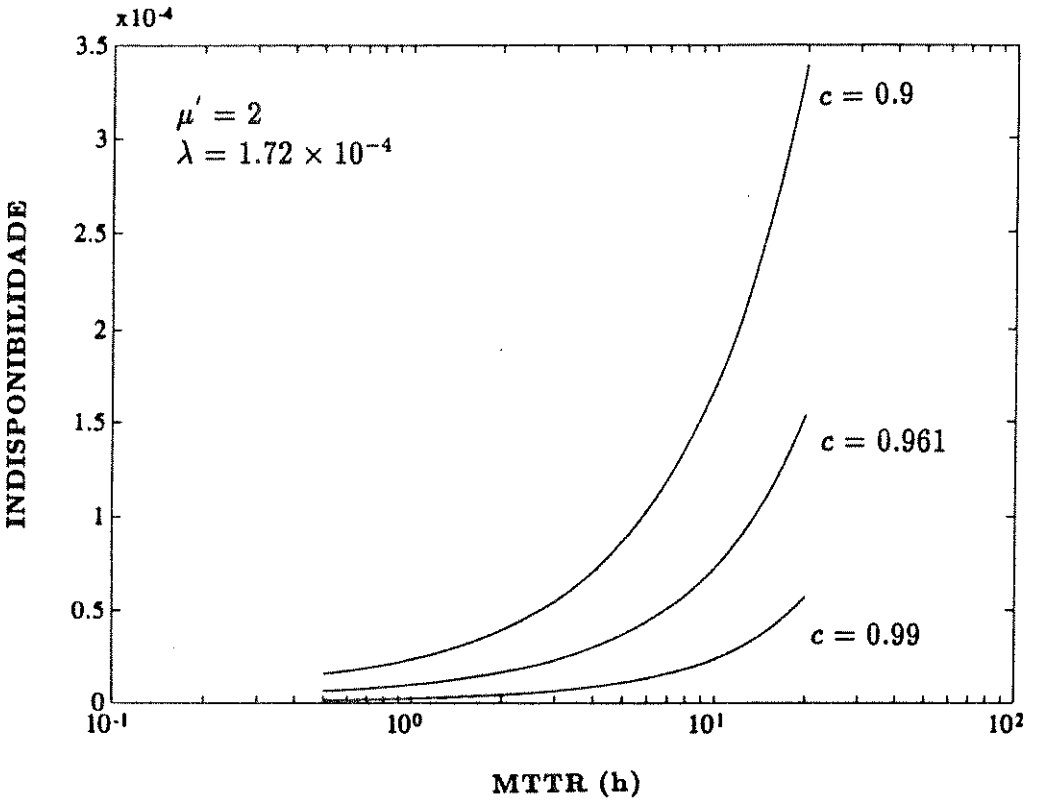


Figura 3.6 - Indisponibilidade do EBR/EBT redundante (parte de modem e controle) versus MTTR.

Com o intuito de obter as probabilidades estacionárias a partir da solução numérica, temos que a matriz de taxas de transição Λ é dada por:

$$\Lambda = \begin{bmatrix} -2\lambda & \lambda(c_A + c_R) & \lambda(1 - c_R) & 0 & \lambda(1 - c_A) \\ \mu & -(\lambda + \mu) & 0 & \lambda & 0 \\ 0 & 0 & -\lambda & \lambda & 0 \\ 0 & \mu & 0 & -\mu & 0 \\ 0 & \mu' & 0 & 0 & -\mu' \end{bmatrix}$$

Aplicando a matriz Λ_m à equação (2.36), e sabendo-se que a disponibilidade é o complemento de I (i.e., $D = 1 - \mathrm{I}$), obtem-se os valores de indisponibilidade apresentados na tabela 3.1, onde é possível notar a influência de c_A e c_R sobre este parâmetro. Para os valores desta tabela, foram considerados $\mu' = 4$, $\mu = 1$ e $\lambda = 1.72 \times 10^{-4}$.

Indisponibilidade	c_A	c_R
35.9×10^{-6}	0.8	0.8
19.6×10^{-6}	0.9	0.9
17.6×10^{-6}	0.95	0.9
16.1×10^{-6}	099	0.9
12.3×10^{-6}	0.9	0.95
6.02×10^{-6}	0.9	0.99
8.12×10^{-6}	0.961	0.961
2.19×10^{-6}	0.99	0.99
0.27×10^{-6}	0.999	0.999
0.059×10^{-6}	1	1

Tabela 3.1 - Indisponibilidade do EBR/EBT redundante (parte de modem e controle) em função de c_A e c_R.

Com relação aos efeitos de c_A e c_R na indisponibilidade dos equipamentos, conforme mostra a tabela 3.1, é interessante notar que a indisponibilidade é mais sensível a variações em c_R do que em c_A. Isto pode ser explicado a partir da simples observação do modelo da figura 3.5, onde uma falha não coberta no plano reserva leva ao estado 3 que, por sua vez, só apresenta transição para um estado de falha. Este aspecto é explorado com maior profundidade no capítulo 4.

3.2.2 Modelo para a ERC

O funcionamento da Estrutura de Referência e Controle compreende a operação conjunta dos EBR's pertencentes à ER0 e ER1. Portanto, no que diz respeito aos processos de falha e reparação, a operação da ERC envolve o funcionamento de quatro planos, onde, na verdade, é necessário que apenas um esteja em condições normais de operação para que o sistema desempenhe suas funções básicas. Deste ponto de vista, para que seja caracterizada falha desta estrutura, é necessário que haja falhas múltiplas (sem reparo) dos quatro planos, ou que ocorra uma falha não coberta no plano ativo da referência ativa com

a consequente não cobertura pela referência reserva (quando há referência reserva).

Baseado nestas características operacionais, é possível identificar dois fatores de cobertura na modelagem desta estrutura: um, c, associado à mudança de planos, abordado em 3.2.1; e outro , c', associado à mudança de referência ativa. Este último representa a probabilidade da ER1 detectar falha da ER0 e assumir a condição de referência ativa, dado que ocorreu uma falha da ER0. Na verdade, esta probabilidade reflete o sucesso da atuação dos mecanismos de supervisão de defeitos e de manutenção de sincronismo empregados no SAMSAT, cujos princípios de funcionamento são descritos em [GON 87] e [FRA 89].

O modelo da figura 3.7 apresenta uma descrição do comportamento operacional da ERC, a partir de um processo markoviano com estado discreto e tempo contínuo. Este modelo possibilita um melhor entendimento sobre o funcionamento da ERC, e, além da hipótese *i*) descrita para o modelo da figura 3.5, é baseado nas seguintes hipóteses:

i) A cobertura é considerada instantânea e relacionada com defeitos hardware permanentes.

ii) Para os EBR's, $c_A = c_R = c$, e uma falha não coberta do plano reserva implica na consideração conservativa de falha deste EBR.

iii) É assumido que a ocorrência de uma falha de sistema gera indicações de alarme suficientes à caracterização da condição de falha. Com isto, o sistema é desativado, iniciando-se os procedimentos de reparo, e como consequência, é possível descartar as probabilidades de ocorrência de falhas nos planos que, por ventura, ainda se encontrem em condições de operação.

iv) Após um reparo, o sistema volta a uma condição operacional.

v) Há um reparador, com taxas μ e μ', para cada um dos dois EBR's, ou seja, o da ER0 e o da ER1. Esta hipótese é justificada pelo fato de que a ER0 e a ER1 estão situadas em pontos geográficos distintos, além do que, estas estações são assistidas localmente.

No modelo da figura 3.7, os estados operacionais são representados por (M,N), onde M e N representam condições associadas ao processo de falha e reparação dos EBR's da ER0 e ER1, respectivamente. M pode assumir os valores de 1 a 6, enquanto N pode assumir apenas os valores de 1 a 4, de tal forma que estes valores representam as condições indicadas a seguir:

1 → dois planos do EBR em condições de operação;

2 → um plano do EBR em falha, mas o equipamento continua em funcionamento;

3 → falha dos dois planos do EBR;

4 → falha não coberta do plano ativo do EBR;

5 → falha do EBR da ER0 (devida a uma falha não coberta pelo equipamento) com a subsequente não cobertura pelo EBR da ER1. Esta situação ocorre quando não há

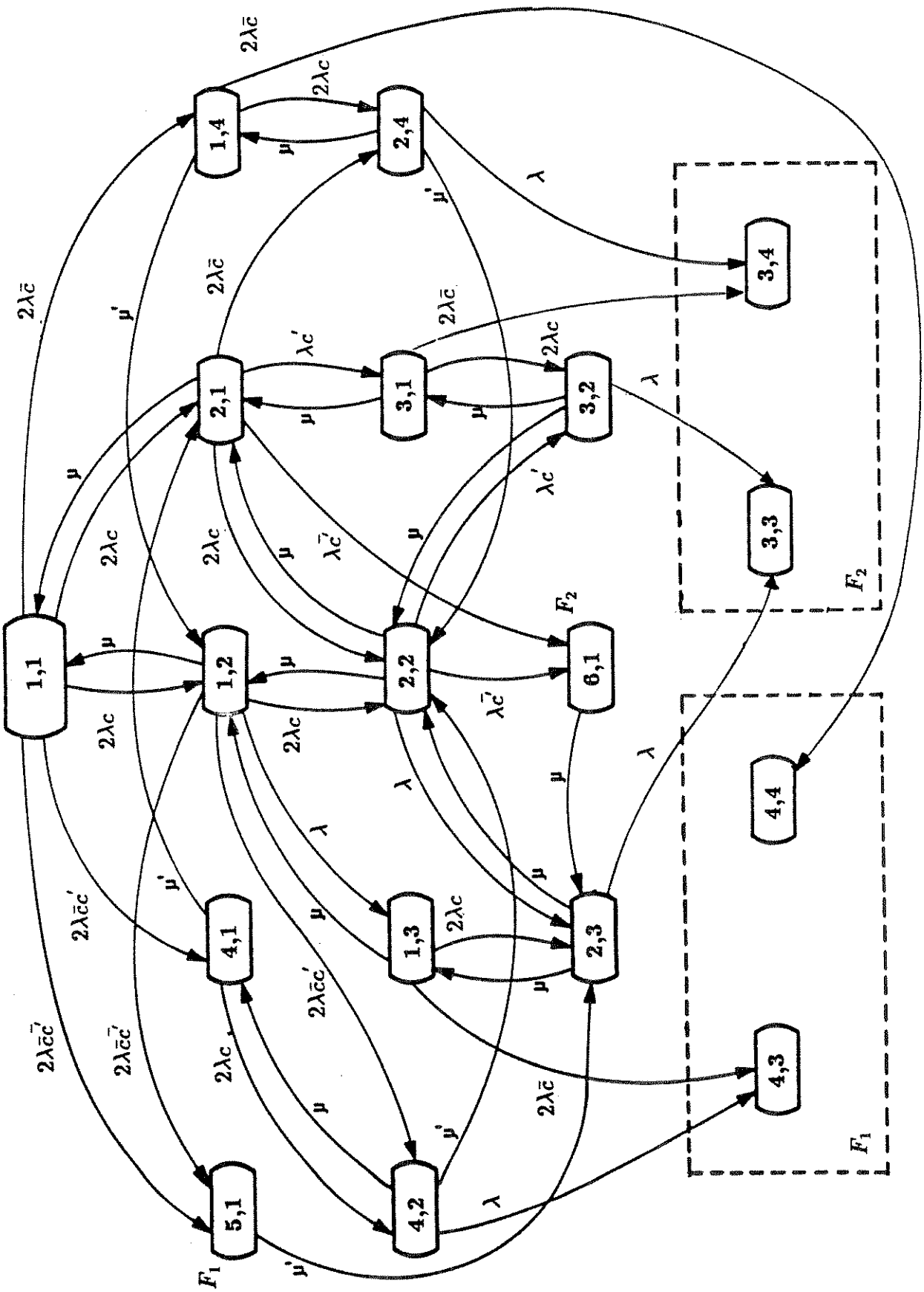


Figura 3.7 - Modelo de Markov para a ERC.

cobertura associada à mudança de referência ativa , e caracteriza falha da ERC;

6 → falha dos dois planos do EBR da ER0 com a subsequente não cobertura pelo EBR da ER1. Esta situação também ocorre quando não há cobertura associada à mudança de referência ativa , e caracteriza falha da ERC.

Por exemplo, o estado constituído por M=1 e N=2, denotado por (1,2), representa a seguinte condição: (dois planos do EBR da ER0 em condições normais de operação; e um plano do EBR da ER1 em falha, mas com este equipamento funcionando). É importante ressaltar que N não pode assumir os valores de 5 e 6, uma vez que tais valores denotam condições estritamente relacionadas com o EBR da ER0.

Os estados em que M e N assumem, simultaneamente, valores maiores ou iguais a 3, ou M=5,6 e N qualquer, representam estados de falha da estrutura. Neste modelo, os seis estados de falha podem ser agrupados em dois estados, F_1 e F_2, com taxas de retorno μ' e μ, respectivamente. Este agrupamento reduz o número de estados do modelo e conduz a resultados ligeiramente mais conservativos, se comparados ao caso onde não há agrupamento. Os resultados conservativos são explicados pelo fato dos estados de falha, quando não agrupados, possuírem, cada um, transição para estados operacionais, e com o agrupamento, estas transições são resumidas a uma só (uma vez que há apenas um reparador para cada estação). Quanto às taxas de transição, este modelo apresenta a mesma notação do modelo da figura 3.5, acrescida de c' que representa o fator de cobertura associado à mudança de referência ativa.

Alguns valores da indisponibilidade da ERC são apresentados na tabela 3.2 em função dos fatores de cobertura c e c'. Estes valores foram obtidos adotando-se a solução numérica dada por (2.36), e considerando-se $c_A = c_R = c$, $\mu' = 4$, $\mu = 1$ e $\lambda = 1.72 \times 10^{-4} falhas/h$.

Indisponibilidade	c	c'
4.33×10^{-7}	0.9	0.95
1.7×10^{-7}	0.961	0.95
0.46×10^{-7}	0.99	0.95
1.33×10^{-7}	0.961	0.961
2.6×10^{-7}	0.9	0.97
1.02×10^{-7}	0.961	0.97
0.276×10^{-7}	0.99	0.97
0.044×10^{-7}	0.999	0.97
4.5×10^{-12}	1	1

Tabela 3.2 - Indisponibilidade da Estrutura de Referência e Controle em função de c e c'.

Como pode ser observado nas tabelas 3.1 e 3.2, a indisponibilidade do EBR/EBT redundante e da ERC é sensível a variações no valor do fator de cobertura. Consequentemente, para que a avaliação propicie resultados precisos, torna-se necessário uma investigação mais

detalhada sobre a atuação dos mecanismos de supervisão, de forma que seja possível determinar o fator de cobertura destes equipamentos.

3.3 Determinação do Fator de Cobertura

A determinação do fator de cobertura de sistemas redundantes é uma tarefa que normalmente acarreta num considerável grau de dificuldade, principalmente pela natureza complexa do processo de detecção & localização de defeitos e de reconfiguração, por envolver dispositivos hardware/software.

Métodos que incluem experimentos de injeção de falhas [ARL 89b] ou o uso de modelagem [TRI 84] e [DUG 89] fornecem resultados com alta precisão, entretanto, a um custo elevado, decorrente do desenvolvimento de estruturas hardware/software voltadas ao desempenho de tais funções. Além do aspecto de custo, o próprio desenvolvimento destas estruturas pode se estender por um período que nem sempre se coaduna com as necessidades de projeto.

Como decorrência destes fatores, na avaliação de segurança de funcionamento do SAMSAT foi elaborada uma metodologia para determinação do fator de cobertura [HOL 89], com o intuito de aproveitar os recursos disponíveis e, consequentemente, atender às particularidades inerentes ao processo de desenvolvimento deste sistema. Esta metodologia é baseada na representação probabilística do fator de cobertura, tendo como dados de entrada as informações coletadas a partir da composição de dois processos: uma componente teórica, a amostragem de dados de projeto, e uma componente de simulação de defeitos. As diretrizes básicas desta metodologia são apresentadas a seguir.

3.3.1 Representação estatística de c

A capacidade de cobertura é indicada por um fator que representa a proporção entre os defeitos cobertos e o total de defeitos[1] aos quais o sistema está sujeito. Desta forma, o fator de cobertura pode ser determinado através de uma avaliação estatística sobre o desempenho dos mecanismos de supervisão face à ocorrência de defeitos.

A avaliação estatística envolve procedimentos baseados num levantamento de dados de falha relacionados com as seguintes premissas:

i) Os defeitos ocorrem numa forma independente, e apenas os defeitos hardware são considerados.
ii) São observados apenas os defeitos classificados como permanentes que, quando ativa-

[1] É importante ressaltar que, em função da análise ocorrer a nível de funcionamento do equipamento, o termo "defeito" é adotado para designar um funcionamento incorreto que poderá ou não conduzir a uma falha do equipamento. Do nosso ponto de vista, este termo é mais adequado, mesmo em detrimento ao fato de que, a nível de funcionamento do plano, este defeito constitui uma falha (conforme a premissa *ii* acima).

dos, conduzam ao não funcionamento do plano[2] ou ao seu funcionamento inadequado.

O levantamento dos dados de defeitos proporciona a formação de um universo de amostras, através do qual o fator de cobertura pode ser determinado, tomando-se como base o conhecimento dos seguintes fatores:

- Composição dos dispositivos funcionais (tais como: gerador de relógio de símbolos, seletor de plano, memórias, etc.), a partir dos quais um defeito pode conduzir o plano à condição de falha.
- Taxas de falha associadas a estes dispositivos.
- Atuação dos mecanismos de supervisão, diante da ocorrência de um defeito.

Desta forma, o fator de cobertura pode ser obtido a partir das expressões (3.4) e (3.5).

$$c = \sum_{i=1}^{N-n} P_i \tag{3.4}$$

$$P_i = \sum_{j=1}^{m} \frac{\lambda_{ij}}{\lambda_T} \quad , \tag{3.5}$$

onde

$N \rightarrow$ número total de defeitos amostrados

$n \rightarrow$ número de defeitos não cobertos

$m \rightarrow$ número de componentes dos dispositivos funcionais que, uma vez acometidos por um defeito, podem causar o defeito coberto i

$\lambda_{ij} \rightarrow$ taxa de falha do componente j, relacionado com o defeito coberto i

$\lambda_T \rightarrow$ taxa de falha total, considerando-se os dispositivos relacionados com todos os defeitos amostrados (cobertos e não cobertos)

A formulação de c através de (3.4), onde P_i é um indicador da complexidade do dispositivo funcional relacionado com o defeito i, tem como objetivo viabilizar uma maneira sistemática de amostragem de defeitos e atribuir pesos aos defeitos, tanto cobertos como não cobertos, em função de suas probabilidades de ocorrência.

Este procedimento confere uma maior precisão ao fator de cobertura levantado, tendo em vista o fato de que determinados defeitos podem ocorrer com maior ou menor probabilidade, dependendo da complexidade associada ao defeito i. Tal complexidade é quantizada através da taxa de falha dos componentes envolvidos pelos dispositvos funcionais que podem ocasionar o defeito i. A taxa de falha dos componentes, por sua vez, pode ser obtida através de bases de dados de falha de componentes ou através de procedimentos baseados em modelos de falha, como por exemplo, a análise por *stress* da norma MIL HDBK 217

[2]Ou item, no caso de um sistema redundante genérico

[MIL 86].

O levantamento de todas estas informações, utilizadas para caracterizar o fator de cobertura, é alcançado a partir da definição do universo de amostras de defeitos.

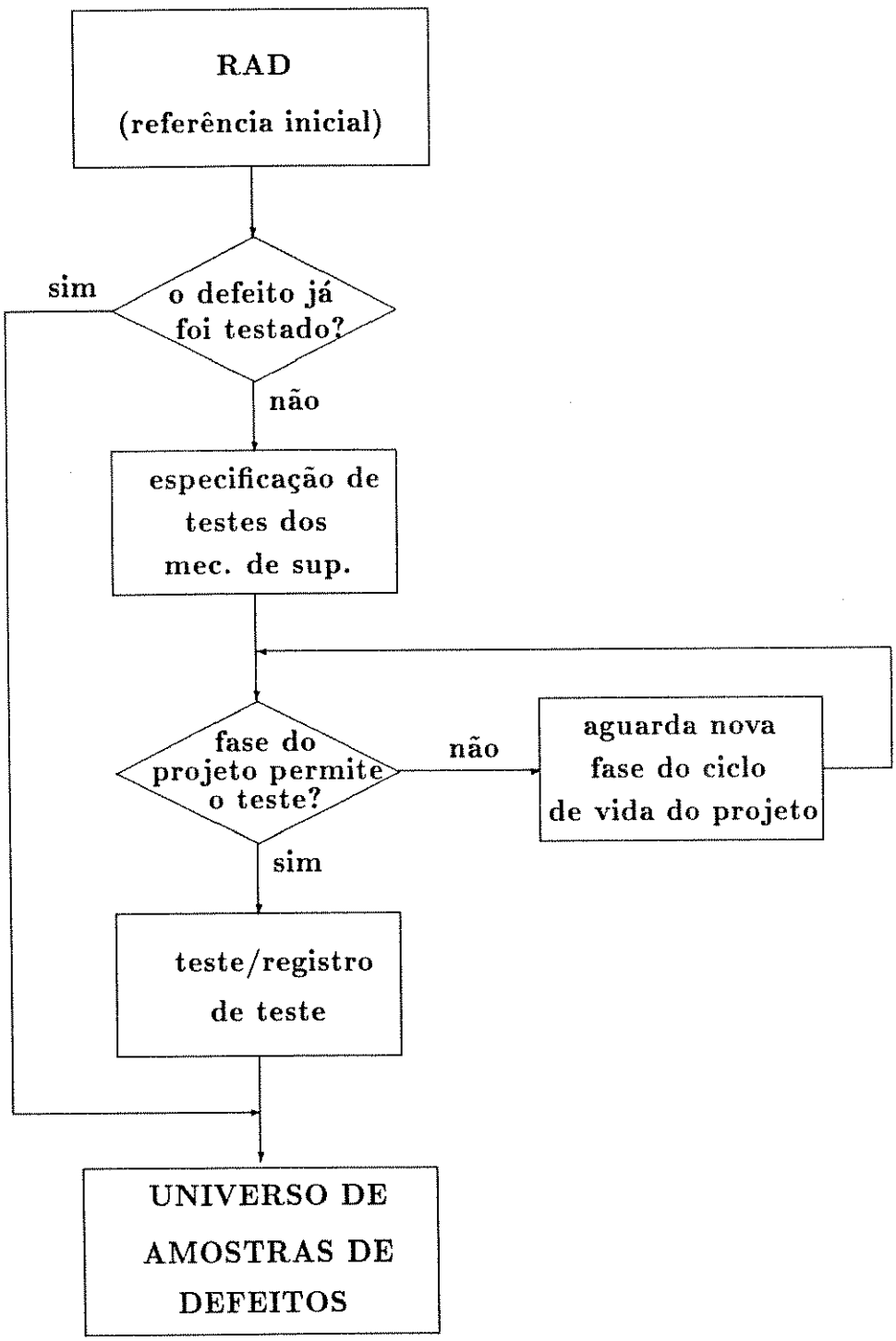


Figura 3.8 - Fluxo do processo de amostragem de defeitos.

3.3.2 Amostragem de defeitos

A obtenção do universo de amostras de defeitos é proporcionada por um processo, cujo fluxo é descrito pela figura 3.8. Este processo é baseado em dois mecanismos básicos:

- Relatórios de amostragem de defeitos, e
- Simulação de defeitos.

A. Relatórios de Amostragem de Defeitos (RAD)

Através do RAD, cujo formato é mostrado na figura 3.9, os projetistas das unidades funcionais relacionam, numa etapa inicial, os defeitos que podem ocorrer em suas unidades, cujas consequências levam à falha do plano. Esta coleta de dados proporciona um levantamento inicial dos dados de falhas & defeitos que refletem a atuação dos mecanismos de supervisão, servindo como ponto de referência para a iniciação do processo de amostragem de defeitos, como pode ser observado na figura 3.8. Após este levantamento inicial, os RAD's são preenchidos em função da observação de defeitos (não descritos anteriormente), levantados ao longo da evolução do processo de amostragem e da maturação do produto em análise.

O agrupamento dos componentes em dispositivos funcionais é usado para classificar os defeitos no RAD, no sentido de avaliar a abrangência dos itens envolvidos e de facilitar a identificação dos mesmos. Para cada defeito são descritas as implicações para o funcionamento do plano, dado que o defeito ocorra, e são relacionadas as circunstâncias mais prováveis para o seu ocasionamento, ou seja, a descrição dos componentes hardware que, submetidos à ocorrência de falhas, podem levar o plano à condição de falha.

A relação das circunstâncias que levam à ocorrência de falha do plano (defeito a nível de equipamento) possibilita a determinação do grau de complexidade dos dispositivos funcionais, através da identificação dos componentes envolvidos e, por extensão, do conhecimento da taxa de falha do dispositivo. Além disso, os relatórios incluem a avaliação da atuação dos mecanismos de supervisão, no que diz respeito à realização dos processos de detecção & localização de defeitos e de reconfiguração dos planos, frente à ocorrência do defeito descrito, através da classificação em defeitos cobertos e não cobertos.

RELATÓRIO DE AMOSTRAGEM DE DEFEITOS (RAD)	No.:
DESCRIÇÃO DO DISPOSITIVO:	
DESCRIÇÃO DO DEFEITO:	
RELAÇÃO DAS CIRCUNSTÂNCIAS MAIS PROVÁVEIS PARA A OCORRÊNCIA DO DEFEITO:	
QUALIFICAÇÕES: O DEFEITO JÁ FOI TESTADO? ☐ SIM ☐ NÃO EM CASO AFIRMATIVO, O DEFEITO É COBERTO? ☐ SIM ☐ NÃO FASE: ☐ DEPURAÇÃO HW ☐ OUTRA ☐ INTEGRAÇÃO HW-SW	

Figura 3.9 - Relatório de Amostragem de Defeitos (RAD): formulário padrão.

B. Simulação de defeitos

A outra componente básica do processo de amostragem de falhas consiste de testes que visam avaliar os aspectos de funcionamento dos mecanismos de supervisão, elaborados a partir da necessidade de complementação ou expansão das informações coletadas nos RAD's.

Estes testes compreendem a simulação de defeitos de certos dispositivos funcionais e a observação do processo de detecção & localização e reconfiguração, e se encontram inseridos em baterias de testes que se adaptam aos recursos disponíveis e às condições proporcionadas pelo estágio de desenvolvimento do produto (e.g., integração hardware-software e

testes de sistema).

Para que possam contribuir para o processo de amostragem, os testes são devidamente registrados através de relatórios apropriados, os quais fazem parte da metodologia de testes do produto em desenvolvimento, e que no caso da amostragem, tem como função facilitar o fluxo de controle dos RAD's.

Todo este processo de amostragem permite a obtenção de um universo de amostras de defeitos, formado com base no funcionamento do sistema segundo condições de contorno impostas pela fase de desenvolvimento em que o projeto se encontra. À medida que o projeto passa por novas fases de seu ciclo de vida, novas estatísticas são levantadas, possibilitando, assim, a ampliação do universo de amostras e a atualização dos resultados obtidos nas fases anteriores. Uma exemplificação deste processo é apresentada a seguir, conforme realizado nos equipamentos redundantes de banda-básica do SAMSAT.

3.3.3 Aplicação da metodologia nos equipamentos SAMSAT

A figura 3.10 ilustra este procedimento com o fator de cobertura do EBR/EBT redundante (parte de modem e controle) ao longo de três fases do ciclo de vida do projeto (CVP) SAMSAT. Estas três fases correspondem, pela ordem cronológica, à: implementação, integração hardware-software, e integração de sistema.

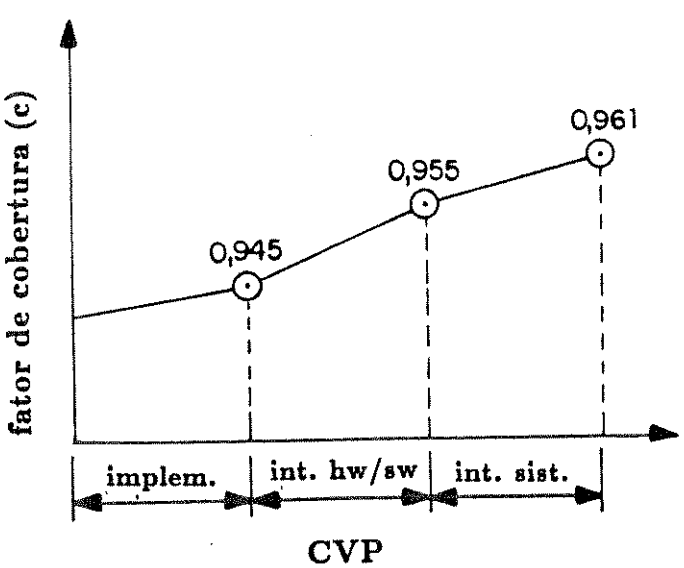


Figura 3.10 - Fator de cobertura do EBR/EBT redundante em função do CVP.

3.4 Estimativas de Indisponibilidade

Esta seção apresenta estimativas para os parâmetros de indisponibilidade sistêmica e de indisponibilidade percebida por um usuário SAMSAT, em termos de falhas de origem hardware. Para a indisponibilidade a nível sistêmico, estes parâmetros correspondem à:

- Indisponibilidade do equipamento de banda-básica das estações ER0 e ER1, i.e., o EBR;
- Indisponibilidade da estrutura redundante de referência e controle (ERC), no que diz respeito aos seus equipamentos de banda-básica.

Para a indisponibilidade percebida pelo usuário, os parâmetros avaliados compreendem:

- Indisponibilidade dos dois tipos de equipamento de banda-básica de usuário, i.e, o EBT redundante e o EBT;
- Indisponibilidade de um enlace SAMSAT, contemplando as configurações possíveis de equipamentos de banda-básica num enlace entre dois usuários.

As estimativas de indisponibilidade são apresentadas em função de alguns valores do MTTR, correspondendo a 0.5, 1 e $2h$ para o EBR e ERC, uma vez que estes equipamentos e esta estrutura são assistidos localmente; e correspondendo a 0.5, 1, 2, 5 e $10h$ para os dois tipos de EBT. Para a indisponibilidade de um enlace, os valores são apresentados através de um gráfico, onde o MTTR varia de 0.5 a $20h$.

3.4.1 Indisponibilidade sistêmica

Conforme mencionado, a aplicação da técnica AMDT depende do desempenho das funções de controle, supervisão e manutenção de sincronismo, as quais são exercidas, de forma centralizada, pela ERC. Consequentemente, a indisponibilidade desta estrutura constitui um parâmetro com informações significativas em termos da segurança de funcionamento do sistema. Baseado neste aspecto, esta subseção apresenta estimativas de indisponibilidade hardware de um EBR e dos equipamentos de banda-básica que integram a ERC, i.e, os EBR's da ER0 e ER1.

A. Indisponibilidade do EBR

A tabela 3.3 apresenta valores de indisponibilidade do EBR (I_{EBR}), os quais foram obtidos a partir do modelo da figura 3.5, considerando-se o fator de cobertura levantado até a fase de integração de sistema, ou seja, $c_A = c_R = 0.961$, conforme mostra a figura 3.10.

I_{EBR}	$MTTR(1/\mu)$
4.86×10^{-6}	$0.5h$
8.12×10^{-6}	$1h$
14.7×10^{-6}	$2h$

Tabela 3.3 - Indisponibilidade hardware do EBR em função do MTTR: valores para $\mu' = 4$, $\lambda = 1.72 \times 10^{-4} falhas/h$ e $c_A = c_R = 0.961$.

B. Indisponibilidade da ERC

A tabela 3.4 apresenta valores de indisponibilidade hardware dos equipamentos de banda-básica da Estrutura de Referência e Controle (I_{ERC}) do SAMSAT. Estes valores foram obtidos a partir do modelo da figura 3.7, e são apresentados em função de $1/\mu$ e $1/\mu'$, onde estes parâmetros refletem a estratégia de reparo empregada. É importante notar que, em decorrência às características de redundância da ERC, a indisponibilidade desta estrutura praticamente só apresenta sensibilidade significativa para variações em μ', ou seja, no MTTR relacionado a uma reconfiguração manual. Este aspecto é compreensível em termos de sistemas com muitas unidades redundantes, porém que apresentam cobertura imperfeita.

I_{ERC}	$MTTR$	
(10^{-7})	$1/\mu$	$1/\mu'$
1.31	$0.5h$	$15min.$
1.33	$1h$	$15min.$
1.4	$2h$	$15min.$
2.64	$1h$	$30min.$
0.896	$1h$	$10min.$

Tabela 3.4 - Indisponibilidade hardware da ERC em função do MTTR: valores para $\lambda = 1.72 \times 10^{-4} falhas/h$ e $c_A = c_R = c' = 0.961$.

No caso do fator de cobertura associado à mudança de referência ativa, c', foi adotado o valor conservativo de $c' = c$, uma vez que pelas características de projeto $c' > c$. A consideração de um valor conservativo para c' se deve ao fato deste parâmetro não ter sido devidamente caracterizado, uma vez que o nível de integração do sistema nas fases anteriores à entrada do sistema em campo não favorecia a realização de determinados testes necessários ao levantamento deste fator de cobertura. Numa rápida explicação, um dos motivos que impossibilitaram a realização destes testes antes da fase de campo foi, por exemplo, a necessidade de posições geográficas distintas para ER0 e ER1.

3.4.2 Indisponibilidade percebida pelo usuário

Do ponto de vista da indisponibilidade percebida por um usuário SAMSAT, dois parâmetros apresentam informações valiosas: a indisponibilidade do equipamento de banda-básica, ao

qual o usuário está conectado; e a indisponibilidade de um enlace entre dois usuários. Esta subseção apresenta estimativas de indisponibilidade hardware do EBT, do EBT redundante e dos equipamentos de banda-básica configurados num enlace SAMSAT. Todos os resultados apresentados dizem respeito à indisponibilidade do ponto de vista de um usuário.

A. Indisponibilidade do EBT

No caso de um EBT não redundante, sua indisponibilidade (I_{EBT}) é expressa a partir de um modelo com dois estados: um, operacional, e, outro, de falha. Assim, I_{EBT} é dada por:

$$I_{EBT} = \frac{MTTR}{MTBF + MTTR} = \frac{\lambda}{\lambda + \mu} \quad , \tag{3.6}$$

onde λ é a taxa de falha do equipamento.

A tabela 3.5 apresenta valores de indisponibilidade hardware do EBT, do ponto de vista de um usuário conectado a uma ULV. A indisponibilidade apresentada nesta tabela foi obtida a partir dos valores de disponibilidade do EBT reportados em [HOL 88].

I_{EBT}	$MTTR(1/\mu)$
8.6×10^{-5}	$0.5h$
17×10^{-5}	$1h$
34×10^{-5}	$2h$
86×10^{-5}	$5h$
172×10^{-5}	$10h$

Tabela 3.5 - Indisponibilidade hardware do EBT em função do MTTR.

B. Indisponibilidade do EBT redundante

Como pode ser visto na figura 3.3, a arquitetura hardware do EBT redundante é composta pela parte de modem e controle, redundante e idêntica ao EBR, mais a parte de interface de linha, composta por ULV/ULG, a qual não apresenta redundância.

Desta forma, do ponto de vista de um usuário conectado a este EBT, a indisponibilidade deste equipamento (I_{EBT-R}) é dada pela seguinte relação:

$$I_{EBT-R} = 1 - D_{EBT-R} = 1 - (1 - I_{EBR})(1 - I_{UL}) \quad , \tag{3.7}$$

onde D_{EBT-R} é a disponibilidade do EBT redundante, I_{EBR} é a indisponibilidade do EBR e da parte de modem e controle do EBT redundante, e I_{UL} é a indisponibilidade da unidade de linha à qual o usuário está conectado, que por sua vez é dada por:

$$I_{UL} = \frac{\lambda_{UL}}{\lambda_{UL} + \mu}$$

onde λ_{UL} é a taxa de falha da unidade de linha configurada.

A tabela 3.6 apresenta estimativas de indisponibilidade hardware do EBT redundante, do ponto de vista de um usuário. Estas estimativas foram obtidas para os valores del_{EBR} dados pela tabela 3.3, I_{UL} para o caso de um usuário conectado a uma ULV, i.e., $\lambda_{UL} = 18.8 \times 10^{-6} falhas/h$, e supondo dois reparadores: um para a parte de modem e controle, e outro para a parte de interface de linha.

I_{EBT-R}	$MTTR(1/\mu)$
1.43×10^{-5}	$0.5h$
2.7×10^{-5}	$1h$
5.23×10^{-5}	$2h$
12.9×10^{-5}	$5h$
26×10^{-5}	$10h$

Tabela 3.6 - Indisponibilidade hardware do EBT redundante em em função do MTTR: valores para $\mu' = 4$ e $c_A = c_R = 0.961$.

A partir dos resultados apresentados, é possível notar que:

- Estes valores correspondem à indisponibilidade de apenas um usuário, e que o fato de um usuário perceber indisponibilidade do EBT redundante não significa, necessariamente, que os outros usuários conectados a este equipamento também percebam. Na verdade, todos os usuários percebem simultaneamente a indisponibilidade do equipamento apenas diante da indisponibilidade da parte modem e controle, cujas estimativas são equivalentes às apresentadas na tabela 3.3, ou diante da probabilidade condicional de que todas as UL configuradas se encontrem indisponíveis (ou seja, uma probabilidade muito pequena).

- O ganho em termos de disponibilidade, decorrente da redundância da parte de modem e controle, não é significativo do ponto de vista de um usuário, como pode ser constatado a partir dos resultados das tabelas 3.5 e 3.6. Entretanto, no que se refere à indisponibilidade percebida simultaneamente por todo o grupo de usuários conectado a um EBT, este ganho é bastante significativo, uma vez que, para o EBT, esta indisponibilidade é praticamente igual à da tabela 3.5, ao passo que, para o EBT redundante, esta indisponibilidade é dada pela tabela 3.3. Tomando como exemplo os valores destas tabelas correspondentes ao MTTR de $1h$, a indisponibilidade percebida por todo o grupo de usuários conectado a um EBT redundante é 20 vezes menor do que a indisponibilidade percebida pelo grupo conectado a um EBT.

C. Indisponibilidade de um enlace

Um enlace entre dois usuários SAMSAT, conforme mostra a figura 3.11, envolve as ETS's às quais os usuários A e B estão conectados, e a ERC. Desta forma, a indisponibilidade de um enlace (I_{enl}), devida aos equipamentos das estações terrestres, é expressa por:

$$\mathrm{I}_{enl} = 1 - (D_{ETS_A} D_{ERC} D_{ETS_B}) \ , \tag{3.8}$$

onde D_{ETS_A} e D_{ETS_B} representam, respectivamente, a disponibilidade das ETS's às quais os usuários A e B estão conectados, e D_{ERC} é a disponibilidade da ERC. É importante notar que a disponibilidade das ETS, bem como da ERC, é dada pela disponibilidade do equipamento de banda-básica e a disponibilidade dos equipamentos de RF, ou seja,

$$D_{ETS} = D_{EBT} D_{ERF} \quad \text{e} \quad D_{ERC} = D_{EBR^*} D_{ERF^*} \quad ,$$

onde D_{EBT} e D_{ERF} representam a disponibilidade do EBT (redundante ou não) e ERF da ETS, respectivamente; e D_{EBR^*} e D_{ERF^*} representam, respectivamente, a disponibilidade dos EBR's e ERF's que integram a ERC.

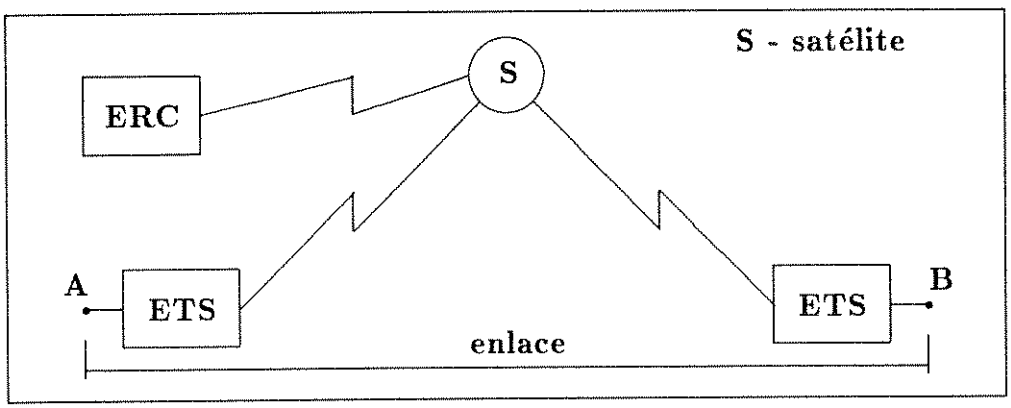


Figura 3.11 - Enlace SAMSAT.

Um aspecto importante relacionado com a equação (3.8) é que a indisponibilidade provocada pelos efeitos de propagação, bem como a indisponibilidade dos equipamentos do satélite, não são incluídas no cômputo da indisponibilidade de um enlace, representada por esta expressão. Isto se deve ao fato de que a consideração destes parâmetros foge ao escopo desta análise. Entretanto, dados referentes aos mesmos podem ser encontrados no relatório 706 do CCIR [CCI 86]].

As curvas da figura 3.12 mostram estimativas da indisponibilidade hardware dos equipamentos de banda-básica envolvidos num enlace, em função do MTTR. A curva 1 corresponde à configuração onde os usuários são conectados a EBT's redundantes. A curva 2 corresponde à configuração em que, numa ponta do enlace, o usuário é conectado a um EBT redundante e, na outra ponta, o usuário é conectado a um EBT não redundante. Já a curva 3 está relacionada com a configuração em que os usuários são conectados a equipamentos não redundantes. Estas curvas implicam, portanto, nas configurações possíveis dos equipamentos de banda-básica envolvidos num enlace SAMSAT.

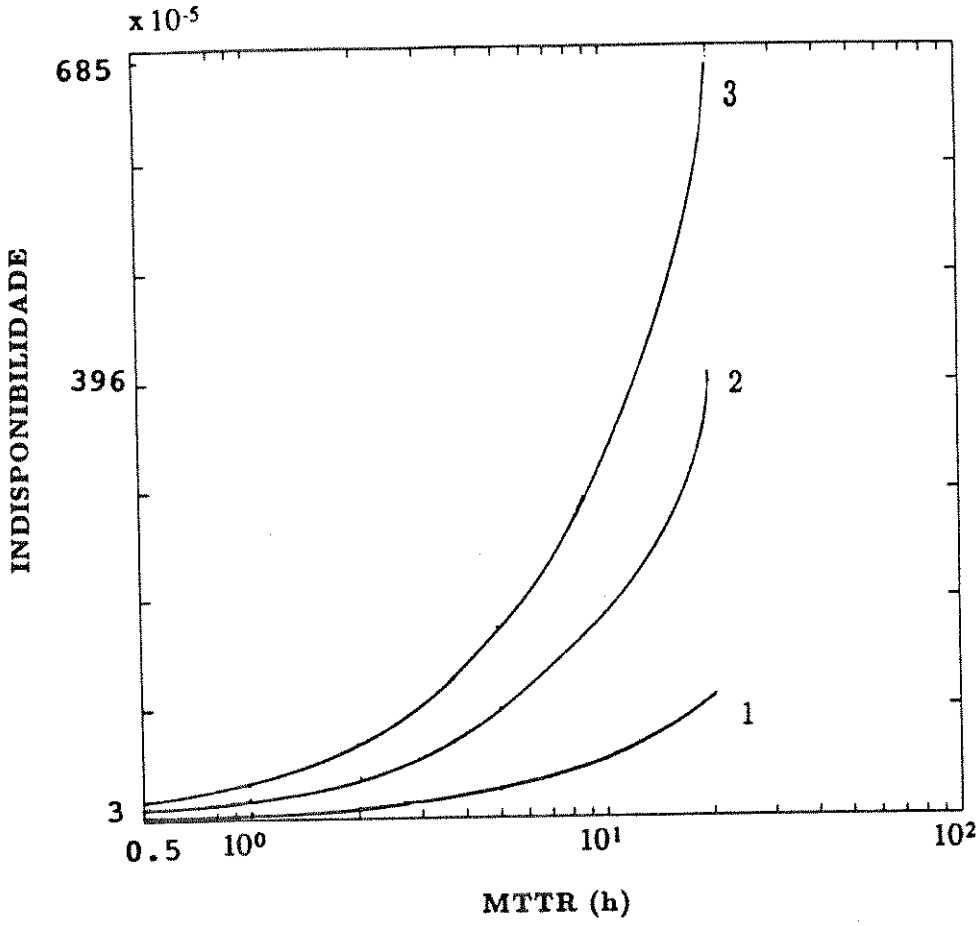


Figura 3.12 - Indisponibilidade hardware de um enlace entre dois usuários SAMSAT.

Através das informações fornecidas pelas curvas, é possível estimar a indisponibilidade hardware de um enlace, devida aos equipamentos de banda-básica envolvidos. Para estimar a indisponibilidade total do enlace, com o propósito de verificar o atendimento ao requisito de indisponibilidade máxima de 0.2% do período de um ano, é necessário considerar a indisponibilidade dos ERF e a indisponibilidade dos equipamentos do satélite. Desta forma, os valores apresentados pelas curvas servem como diretrizes em termos de escolha da configuração de equipamentos de um enlace, bem como na determinação de uma estratégia de reparo que forneça o suporte de manutenção adequado, de tal forma a atender ao requisito de segurança de funcionamento e às necessidades do usuário.

3.5 Conclusões

Este capítulo apresentou o processo de avaliação da segurança de funcionamento do SAMSAT, ilustrando a aplicação de algumas técnicas descritas no capítulo 2, e, além disso, propondo uma metodologia para levantamento do fator de cobertura dos equipamentos redundantes. Calcado no que foi apresentado, é possível ressaltar os seguintes aspectos:

- Um elevado nível de segurança de funcionamento é essencial para determinados equipamentos de sistemas de comunicação de dados com controle e operação centralizados, como o apresentado pelo SAMSAT. Desta forma, a garantia deste nível está atrelada à eficiência de utilização dos mecanismos de tolerância a defeitos, que possibilita a operação confiável dos equipamentos essenciais e, por extensão, do sistema como um todo.

- A utilização eficiente das estruturas redundantes está diretamente relacionada com uma alta capacidade de cobertura, de forma que fatores de cobertura adequados constituem aspectos de segurança de funcionamento que devem ser perseguidos durante a especificação e desenvolvimento de tais estruturas. Neste particular, a análise da cobertura assume um papel relevante no que se refere à avaliação dos parâmetros de segurança de funcionamento, ou mesmo à obtenção de indicativos sobre a atuação dos mecanismos de supervisão.

- A metodologia para determinação do fator de cobertura, apresentada na seção 3.3. demonstrou-se eficiente ao longo de sua aplicação no SAMSAT, principalmente no tocante à rapidez de resposta e à utilização dos recursos disponíveis durante o desenvolvimento do produto. Estes fatores caracterizam a metodologia como uma alternativa prática e não onerosa.

- A avaliação de indisponibilidade dos equipamentos, estrutura de controle, e enlace, conforme reportada neste trabalho, proporciona a constatação de outro aspecto importante: a influência decisiva da estratégia de reparo na obtenção de uma alta disponibilidade, sobretudo para os equipamentos não redundantes. Consequentemente, a continuidade do serviço oferecido aos usuários de um sistema como o SAMSAT é determinada não só pelo uso de mecanismos de tolerância a defeitos mas também por medidas gerenciais voltadas à operação e manutenção do sistema.

Capítulo 4

Avaliação do PSN COMPAC

O estudo de caso apresentado neste capítulo corresponde à avaliação de segurança de funcionamento de um nó de comutação de pacotes (PSN) do sistema COMPAC. Este sistema [ARR 87] foi especificado e desenvolvido pelo CPqD-TELEBRÁS com o propósito de constituir redes de comunicação de dados. Neste sentido, a finalidade básica do sistema COMPAC é oferecer serviços de comutação de pacotes, bem como facilidades de supervisão e controle que possam auxiliar a administração da rede.

A falha de um PSN acarreta algumas consequências indesejáveis, tais como, isolamento dos usuários conectado a este nó, alteração da topologia da rede, perda de tarifação, etc. Estes problemas degradam o desempenho de alguns parâmetros, e.g., velocidade e confiabilidade, associados à qualidade de serviço percebida pelos usuários e pela gerência da rede.

Em função da natureza distribuída de um PSN, a segurança de funcionamento de toda a estrutura hardware do nó assume um papel importante. Isto pode ser melhor compreendido através do fato de que o comprometimento das principais funções do nó pode ser proporcionado a partir de falhas hardware ocorridas em qualquer parte de um enlace entre dois usuários, ou seja, a falha pode estar localizada tanto nos circuitos de acesso ao nó, como na própria estrutura de controle, a qual é comum a todos os usuários do referido nó.

Assim, a segurança de funcionamento de um PSN pode ser dividida em três níveis associados à sua estrutura:

- A segurança de funcionamento da estrutura de controle do nó;
- A segurança de funcionamento dos circuitos de acesso ao nó;
- A segurança de funcionamento percebida pelo usuário.

No que diz respeito aos sistemas redundantes e reparáveis, utilizados em aplicações com tolerância a pequenos intervalos de interrupção do serviço, onde se inclui o PSN do sistema COMPAC, a indisponibilidade se apresenta como o indicativo de segurança de funcionamento que melhor traduz o nível de satisfação de seus usuários. Desta maneira, a avaliação de segurança de funcionamento do PSN COMPAC, apresentada neste relatório, é baseada na análise dos seguintes parâmetros:

- A indisponibilidade de duas das principais funções providas pela estrutura de controle do nó (designada como estrutura MA/SA):
  - Indisponibilidade das funções de estabelecimento de chamada;
  - Indisponibilidade das funções de tarifação.
- A indisponibilidade da interface de usuário (LA);
- A indisponibilidade percebida pelo usuário.

Em termos de funcionamento da estrutura MA/SA, a indisponibilidade das funções de estabelecimento de chamada, por um lado, constitui um importante indicativo a ser considerado no cômputo da indisponibilidade percebida pelo usuário, uma vez que representa a probabilidade de um usuário não dispor do hardware, compartilhado por todos os usuários do nó, o qual é responsável pelo fornecimento dos circuitos virtuais necessários ao estabelecimento de uma chamada. Por outro lado, a indisponibilidade destas funções leva à perda de receita, dado que existe uma demanda pelos serviços, e os recursos não são fornecidos, caracterizando, desta maneira, entraves à operação e gerência da rede. Neste aspecto, a indisponibilidade das funções de tarifação fornece um outro indicativo também valioso à gerência de rede, justamente pelo fato de que, em tal situação, o serviço fornecido não está sendo tarifado.

Através da indisponibilidade da interface de usuário (LA) é possível avaliar as chances de um determinado usuário não conseguir acesso ao nó ou perder a comunicação estabelecida, em função de uma falha em seus circuitos de acesso. Este parâmetro representa um papel importante no tocante à indisponibilidade percebida por um usuário, conforme apresentado posteriormente.

Nesta avaliação, os valores de indisponibilidade são apresentados em função do MTTR, de modo a evidenciar o impacto da estratégia de reparo na indisponibilidade do PSN, e da cobertura de defeitos associada às estruturas redundantes (no caso, a estrutura MA/SA), com o propósito de fornecer uma análise baseada em princípios de funcionamento mais realistas.

Este capítulo está estruturado da seguinte maneira: A seção 4.1 apresenta uma breve descrição da arquitetura e das características operacionais de um PSN COMPAC, com o objetivo de facilitar o entendimento dos modelos de segurança de funcionamento; A seção 4.2 traz os modelos de segurança de funcionamento utilizados na avaliação; A seção 4.3 mostra os efeitos da cobertura na indisponibilidade da estrutura MA/SA; A seção 4.4 apresenta estimativas de indisponibilidade dos parâmetros avaliados, destacando o impacto significativo de um procedimento de comutação periódica, aplicado entre os módulos ativo e reserva da estrutura MA/SA, sobre estes parâmetros; Finalmente, a seção 4.5 apresenta alguns comentários sobre os resultados obtidos.

4.1 Descrição do PSN COMPAC

O PSN COMPAC é responsável pelo desempenho das seguintes funções básicas [ARR 87]:

- Estabelecimento, manutenção e desconexão de circuitos virtuais.
- Coleta de dados para tarifação.
- Coleta de dados para a produção de estatísticas.
- Supervisão interna do PSN, tanto a nível hardware quanto a nível de software.
- Comunicação com os outros elementos da rede.

A arquitetura hardware e as características operacionais deste PSN são descritas nas subseções a seguir.

4.1.1 Arquitetura hardware

A estrutura hardware do PSN é modular, sendo a estação a unidade básica para a sua composição. Uma estação contém um conjunto de placas de circuito impresso, interligadas e configuradas para desempenhar uma ou mais funções do nó.

O PSN é constituído por três tipos de estação:

- Estação de disco (DST) : a qual comporta o módulo MA.
- Estação básica (BST) : a qual comporta o módulo SA.
- Estação de linha (LST) : a qual comporta o módulo LA.

Com o objetivo de simplificar a nomenclatura utilizada neste relatório para designar as partes hardware, os três tipos de estação passam a ser representados pelo módulo correspondente, uma vez que proporciona a associação direta entre uma unidade hardware e suas características funcionais. Desta forma, a arquitetura hardware de um PSN COMPAC é constituída pelos módulos MA, SA, LA, segundo algumas configurações possíveis. A figura 4.1 apresenta uma destas configurações.

Em linhas gerais, o módulo MA é responsável pelas funções de gerenciamento do sistema, sendo, inclusive, configurado com uma unidade de memória de massa. O módulo SA também desempenha funções de gerenciamento, sendo responsável, entre outras, pelas funções de estabelecimento de chamada. O módulo LA tem como função proporcionar a interface entre o PSN e os assinantes e troncos.

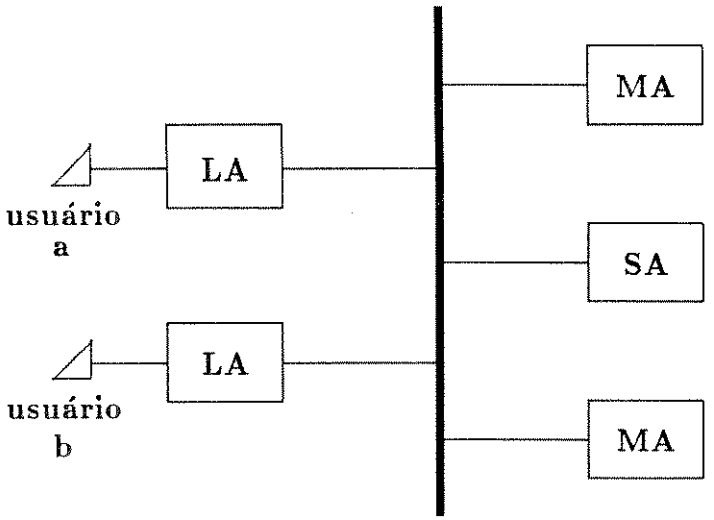


Figura 4.1 - Arquitetura de um PSN COMPAC:
configuração usual.

4.1.2 Características operacionais

Do ponto de vista funcional, o PSN COMPAC pode ser dividido segundo dois grupos de funções: um, representado pela estrutura composta pelos módulos MA e SA (estrutura MA/SA), a qual é responsável pelas funções de controle e gerenciamento do nó; e outro, caracterizado pelas funções de interface desempenhadas pelo módulo LA. As características associadas ao modo de operação da estrutura MA/SA e do módulo LA são descritas a seguir.

A. Estrutura MA/SA

A estrutura MA/SA possui flexibilidade de composição de sua arquitetura, apresentando algumas opções de configuração. Uma configuração típica corresponde à utilização de dois MA's, onde, durante funcionamento normal, um se encontra como ativo (MA_a) e o outro na condição reserva (MA_r), podendo o ativo passar a reserva e vice-versa. Além disso, o MA_r pode assumir as funções do SA, quando este é acometido por falha. Para efeito de análise, esta configuração é representada da seguinte forma (MA_a, MA_r, SA).

Em termos de funcionamento da estrutura, a não existência de um MA em condições operacionais leva, por exemplo, à perda das funções de tarifação. Por outro lado, a falha do SA, sem a subsequente substituição pelo MA_r, implica na perda das funções de estabelecimento de chamada.

Outra possibilidade de composição consiste em configurar a estrutura MA/SA com mais de um SA, além dos dois MA's. Neste tipo de configuração, os usuários são distribuídos entre nSA's, enquanto um SA é colocado como reserva, de forma a substituir um dos outros SA's que, por ventura, venha a falhar. Para efeito de análise, esta configuração é representada da seguinte forma (MA_a,MA_r,$SA(n+1)$).

O objetivo da configuração com mais de um SA é distribuir o processamento de chamadas entre os SA e , assim, melhorar o desempenho destas funções nos casos onde há um aumento significativo da carga de processamento da estrutura. Normalmente, estes casos são proporcionados por um determinado perfil de tráfego onde as chamadas de curta duração são predominantes e, portanto, as funções de estabelecimento de chamada são mais solicitadas.

B. Módulo LA

Os usuários são agrupados e conectados a um módulo LA através de unidades de interface (LB), as quais podem ser diferenciadas em três tipos (LB2, LB8, LB16), atendendo, respectivamente, a grupos de 2, 8 e 16 usuários. Qualquer falha em uma determinada LB implica apenas na não operacionalidade dos assinantes conectados a esta placa, e os demais assinantes não sofrerão qualquer degradação na qualidade de serviço.

4.2 Modelos de Segurança de Funcionamento

Dentre as funções exercidas pela estrutura MA/SA, é possível destacar duas de particular interesse: funções de estabelecimento de chamada e funções de tarifação. Os modelos que possibilitam a obtenção das medidas de indisponibilidade destas duas funções são descritos nesta seção. Estes modelos descrevem o comportamento operacional da estrutura, no que diz respeito aos processos de falha e reparação vinculados ao fornecimento das funções de estabelecimento de chamada e de tarifação.

Os modelos apresentados contemplam configurações dos tipos (MA_a,MA_r,SA) e (MA_a, MA_r,SA($n+1$)). Para o segundo tipo, a configuração dos módulos SA (configuração SA($n+1$) é modelada em separado, e os resultados obtidos são incorporados ao modelo da estrutura. Tal procedimento é apresentado mais adiante, através da análise realizada para a configuração SA($2+1$).

4.2.1 Modelos para a estrutura MA/SA

O comportamento operacional da estrutura MA/SA no fornecimento das funções de estabelecimento de chamada e de tarifação é representado através de processos markovianos com estado discreto e tempo contínuo, conforme os modelos das figuras 4.2 e 4.3, os quais também são encontrados em [HOL 92] (apenas o primeiro) e [HOL 91c]. Nestes modelos foram adotadas as seguintes hipóteses:

i) O barramento não é considerado em função de apresentar uma taxa de falha desprezível em relação às demais.

ii) A cobertura é considerada instantânea e relacionada com defeitos hardware permanentes.

iii) São consideradas coberturas diferentes para os módulos ativo e reserva , com o intuito de enriquecer a análise dos efeitos da cobertura na indisponibilidade da estrutura.

iv) É assumido que a ocorrência de uma falha do sistema gera indicações de alarme suficientes à caracterização da condição de falha. Com isto, o sistema é desativado, iniciando-se os procedimentos de reparo, e como consequência, é possível descartar as probabilidades de ocorrência de falhas nos módulos que, por ventura, ainda se encontrem em condições de operação.

v) É considerado que, após um reparo, o sistema volta às condições normais de operação, e que há apenas um reparador com taxas μ e μ'.

vi) São consideradas até duas falhas consecutivas sem reparo.

A representação dos estados dos modelos correspondentes às figuras 4.2 e 4.3 é feita a partir de uma notação vetorial da forma (x, y, z). Os elementos do vetor indicam, respectivamente, as condições operacionais dos módulos MA_a, MA_r e SA , podendo assumir três valores, $1, \bar{1}$ e 0, onde:

1 representa condição normal de operação do módulo
$\bar{1}$ representa falha não coberta do módulo
0 representa falha do módulo

Os estados de falha podem ser agrupados segundo as possibilidades de retorno a um estado operacional, como pode ser observado no modelo da figura 4.2, ou seja:

Estado F_1 $\rightarrow$ conjunto de estados de falha decorrentes de defeitos não cobertos pelo sistema. Caracteriza um estado de falha com taxa de transição μ' (decorrente de uma reconfiguração manual) a um estado operacional.

Estado F_2 **e** F_3 $\rightarrow$ conjunto de estados de falha com taxa de transição μ a um estado operacional.

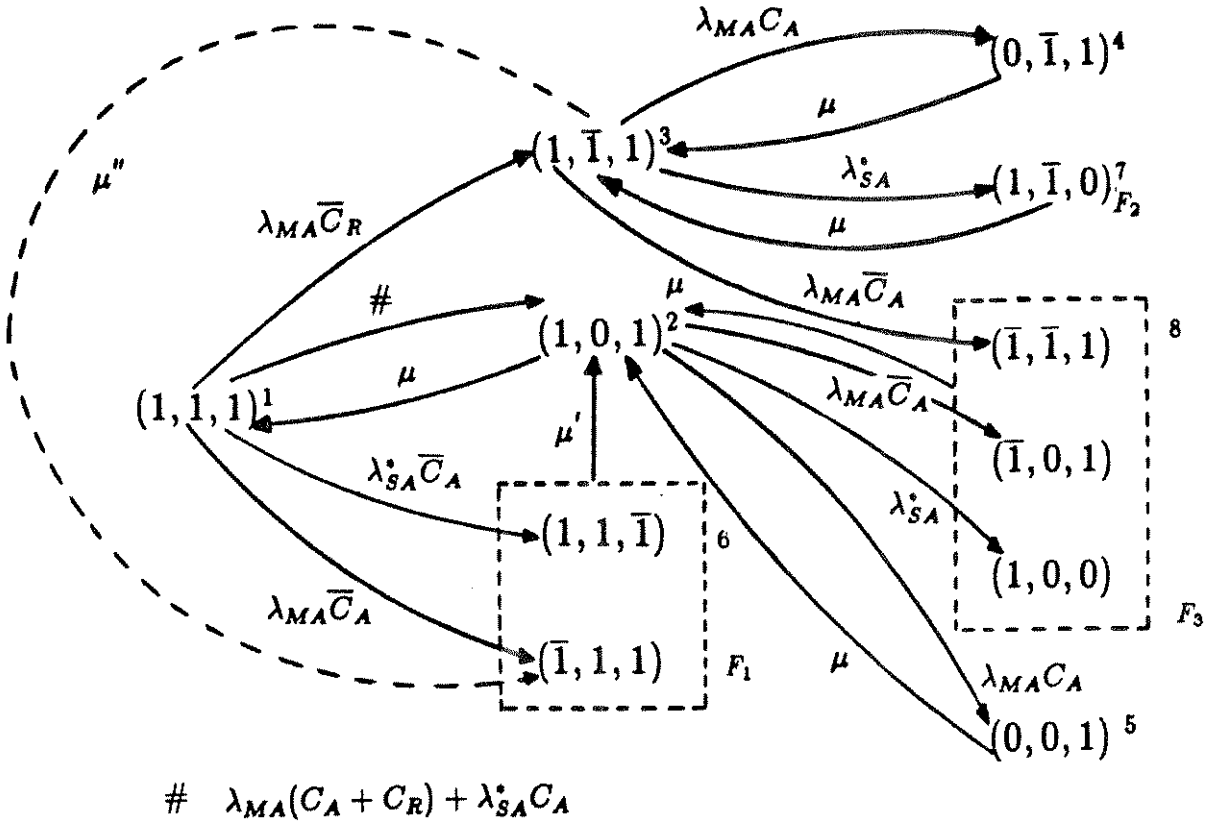


Figura 4.2 - Modelo de Markov para as funções de estabelecimento de chamada da estrutura MA/SA.

No modelo da figura 4.3, conforme pode ser observado, há apenas o agrupamento de estados de falha caracterizado pelo estado F_3. Contudo, é possível notar que as taxas de retorno entre os estados de falha e os estados operacionais também apresentam diferenciação: os estados F_1 e F_4 possuem taxa de transição μ', e os estados F_2 e F_3 apresentam taxa μ.

No tocante às taxas de transição entre os estados, os dois modelos apresentam a seguinte notação:

$\lambda_{MA}, \lambda_{SA} \rightarrow$ taxas de falha dos módulos MA e SA, respectivamente.

$C_A, C_R \rightarrow$ fatores de cobertura associados aos módulos ativo e reserva, respectivamente. É importante lembrar que $\overline{C}_A$ e $\overline{C}_R$ correspondem, respectivamente, a $(1 - C_A)$ e $(1 - C_R)$.

$\mu \rightarrow$ taxa de reparo (falha permanente), $\mu = 1/MTTR$.

$\mu' \rightarrow$ taxa de reparo de uma falha não coberta da unidade ativa (falha permanente). Corresponde a uma reconfiguração manual, $\mu' > \mu$.

$\mu'' \rightarrow$ taxa de transição decorrente de uma comutação periódica entre os MA's, $\mu'' << \mu$, baseada nas mesmas hipóteses e considerações descritas na subseção 2.2.4.

Nos modelos das figuras 4.2 e 4.3, as transições do estado 3 para o estado 6 (fig. 4.2) e do estado 3 para o estado 9 (fig. 4.3) são representados de forma tracejada. O motivo desta representação especial deve-se ao fato de que, para enriquecer a avaliação, foram consideradas duas situações: *i*) existe um procedimento de comutação periódica entre os MA's ativo e reserva, e *ii*) não existe tal procedimento.

i) No caso do procedimento de comutação periódica entre os MA's ser adotado, é possível uma transição do estado 3, caracterizada por uma falha não coberta do MA_r, para o estado 6 (no caso da figura 4.2) ou para o estado 9 (no caso da figura 4.3). Estes dois últimos estados possuem, nos seus respectivos modelos, uma taxa de transição μ' (mais rápida, se comparada à taxa μ) a um estado operacional. No modelo da figura 4.2, a taxa de transição de 3→6 (da mesma forma que a taxa de 3→9, no modelo da figura 4.3), representada por μ'', é determinada pelo período com que a comutação entre os MA's é realizada.

Conforme pode ser observado na análise da subseção 2.2.4 e em [HOL 91b], a realização de supervisões periódicas em unidades reservas proporciona um ganho de disponibilidade significativo, em se tratando de sistemas com características de redundância como as apresentadas pela estrutura MA/SA. Além disso, este ganho de disponibilidade apresenta pouca sensibilidade a variações no período de comutação (se comparado a variações de mesma ordem no MTTR), o que possibilita uma razoável flexibilidade quanto à escolha deste período.

Efeito semelhante pode ser obtido com a realização de um procedimento de comutações periódicas entre os MA's, o qual pode detectar, de forma indireta, a presença de falhas da unidade reserva (decorrentes de defeitos não cobertos), uma vez que esta unidade ao passar para a condição ativa leva a estrutura a um estado de falha com retorno mais rápido (μ') a um estado operacional.

A comutação periódica entre os MA's apresenta a vantagem de ser facilmente implantada, requerendo apenas procedimentos operacionais (e.g., comutações programadas para dia e horário de baixo tráfego). Este aspecto é muito útil, pois permite que o procedimento de comutação periódica seja aplicado, inclusive, a sistemas que se encontram em fases avançadas de seu ciclo de vida (por exemplo, aplicação comercial), sem que haja necessidade de alterações de produto, normalmente onerosas.

ii) No caso do procedimento de comutação não ser adotado, i.e., $\mu'' = 0$, não existem as transições entre os estados 3 e 6 da figura 4.2, e entre os estados 3 e 9 da figura 4.3.

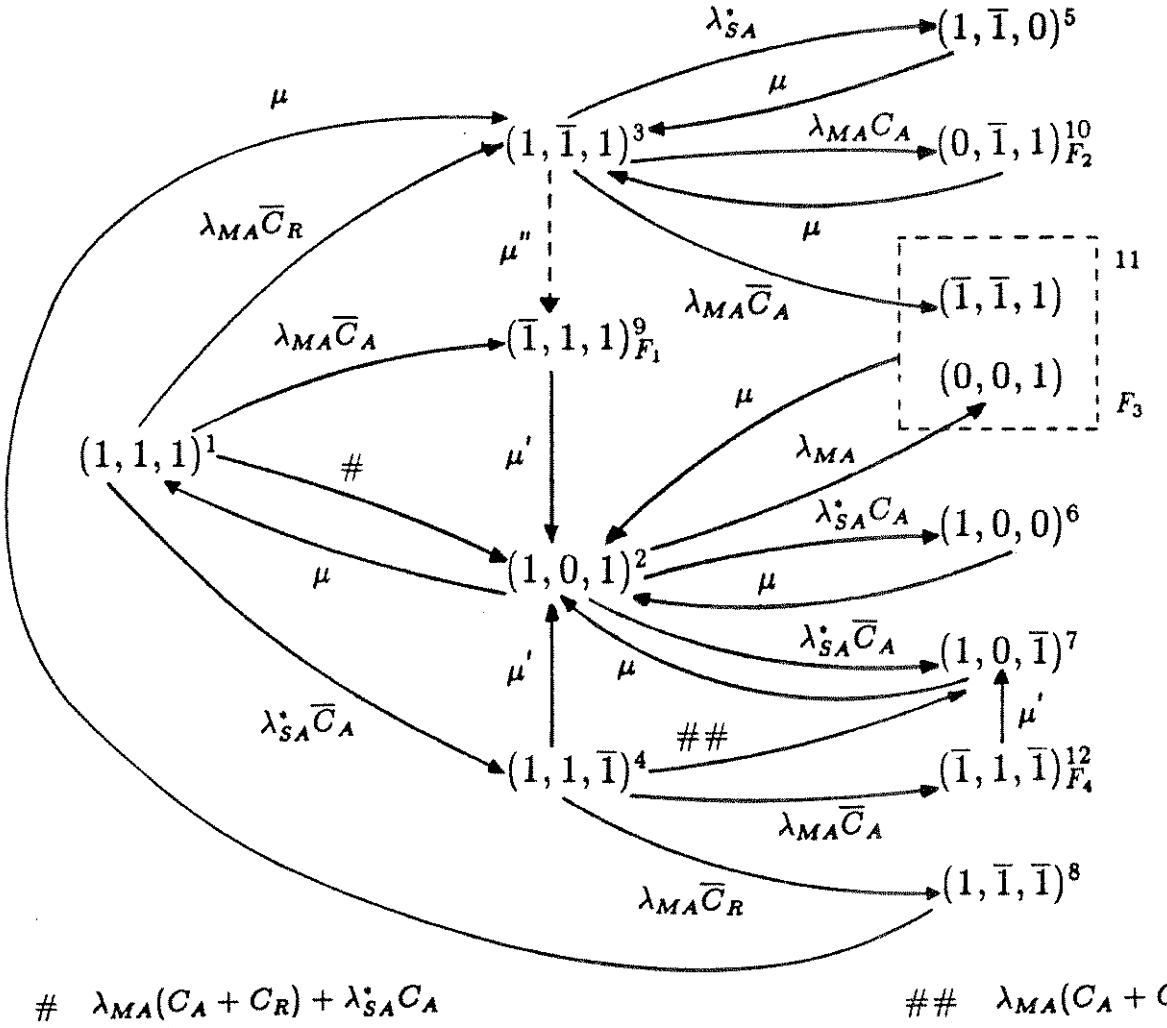


Figura 4.3 - Modelo de Markov para as funções de tarifação da estrutura MA/SA.

Ainda com relação aos modelos das figuras 4.2 e 4.3, a existência do λ_{SA} equivalente (λ^*_{SA}) em algumas taxas de transição reflete a possibilidade do PSN ser configurado com mais de um SA. Nestes casos, a taxa de falha λ^*_{SA} é obtida a partir da avaliação da estrutura SA(n+1) em particular. Caso contrário, quando a estrutura MA/SA é configurada com apenas um SA, o λ^*_{SA} corresponde ao próprio λ_{SA}. A seguir é apresentada a análise da configuração SA(2+1), com o intuito de avaliar o comportamento desta estrutura mediante a ocorrência de falhas e, por extensão, obter o λ^*_{SA} a ser utilizado no modelo da estrutura MA/SA para a configuração (MA_a,MA_r,SA(2+1)).

4.2.2 Configuração SA(2+1)

Na configuração SA(2+1), representada na modelagem por (SA_1, SA_2, SA_r), os assinantes são distribuídos entre dois SA's (SA_1 e SA_2), os quais podem ser substituídos pelo SA reserva (SA_r) em caso de falha.

Em função destas características, a análise apresentada a seguir é conduzida segundo o ponto de vista de ocorrência de falha de SA percebida por um usuário, uma vez que, devido à distribuição dos usuários entre os SA's configurados, o surgimento de falhas pode acarretar indisponibilidade das funções providas pelo SA apenas para um determinado grupo de usuários. Desta forma, os resultados provenientes da análise com este enfoque fornece subsídios à avaliação de indisponibilidade relacionada com o processo de estabelecimento de chamada de um usuário, conforme é abordado na subseção 4.4.4.

Para esta análise, é utilizado o modelo de descrição de estados apresentado na figura 4.4. Este modelo corresponde a um processo markoviano a estado discreto e tempo contínuo, onde foram adotadas as seguintes hipóteses:

i) São consideradas até duas falhas consecutivas sem reparo.

ii) Existe apenas um reparador com taxa μ.

iii) Os estados de falha são considerados como estados absorvedores (sem retorno a um estado operacional), uma vez que o objetivo é determinar o tempo médio de ocupação dos estados operacionais, como será visto a seguir.

iv) É assumido que em t=0 o sistema se encontra com todos os módulos em condições de operação.

v) A cobertura é considerada instantânea e relacionada com defeitos hardware permanentes.

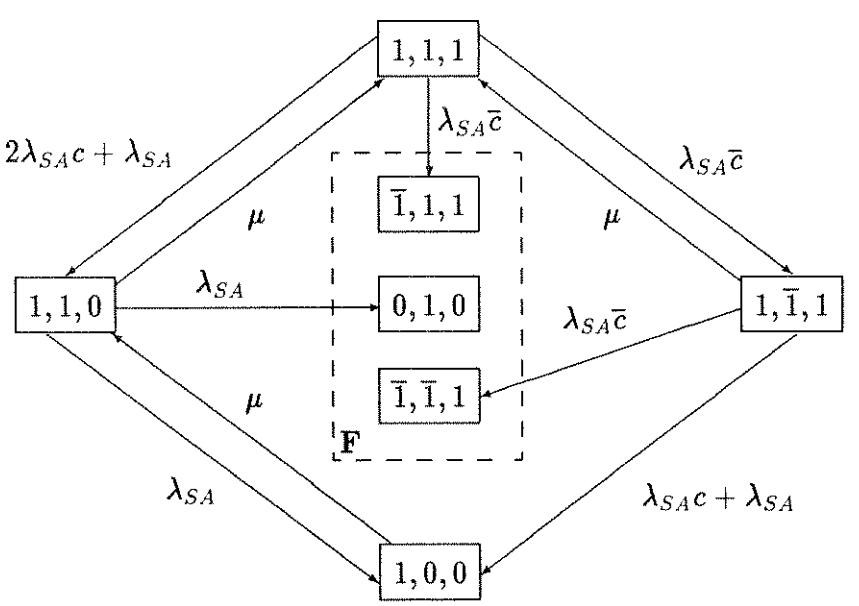


Figura 4.4 - Modelo de Markov para a configuração SA(2+1).

Similarmente à representação adotada nos modelos da estrutura MA/SA, os estados do modelo da figura 4.4 são caracterizados a partir de uma notação vetorial da forma (x, y, z). Os elementos do vetor indicam, respectivamente, as condições operacionais dos módulos SA_1, SA_2 e SA_r , podendo assumir três valores, 1, $\bar{1}$ e 0, onde:

1 representa condição normal de operação do módulo
$\bar{1}$ representa falha não coberta do módulo
0 representa falha do módulo

O modelo da figura 4.4 tem a seguinte notação:

$\lambda_{SA} \rightarrow$ taxa de falha do módulo SA

$c \rightarrow$ fator de cobertura associado à substituição de um SA em falha pelo SA_r

$\mu \rightarrow$ taxa de reparo (falhas permanentes), $\mu = 1/MTTR$

A partir do modelo da figura 4.4 é possível obter a matriz de transição de estados apresentada em (4.1).

$$\Lambda = \begin{pmatrix} -3\lambda_{SA} & \lambda_{SA}(2c+1) & \lambda_{SA}\bar{c} & 0 & \lambda_{SA}\bar{c} \\ \mu & -(2\lambda_{SA}+\mu) & 0 & \lambda_{SA} & \lambda_{SA} \\ \mu & 0 & -(2\lambda_{SA}+\mu) & \lambda_{SA}(c+1) & \lambda_{SA}\bar{c} \\ 0 & \mu & 0 & -\mu & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad (4.1)$$

Nesta análise, o interesse é, sobretudo, determinar uma taxa de falha equivalente (λ^*_{SA}) que simbolize, nos modelos das figuras 4.2 e 4.3, a transição entre um estado com SA operacional e um estado com SA em falha. Uma aproximação razoável para esta determinação é considerar que $\lambda^*_{SA} = 1/MTTF$, onde o MTTF (tempo médio para falhar) representa o tempo médio de ocupação dos estados operacionais do modelo da figura 4.4. Desta forma, utilizando a expressão (2.35) apresentada na subseção 2.2.2, tem-se:

$$MTTF = \frac{P_o[-A_{oo}]^{-1}1_k}{P_o 1_k} \qquad (2.35)$$

onde P_o é o vetor das probabilidades estacionárias de ocupação dos estados operacionais, $P_o = [P_1, ..., P_{k=4}]$, obtido do modelo da fig.4.4 acrescido da taxa de transição μ entre o estado de falha e o estado operacional (1,1,0); Λ_{oo} é a submatriz de tamanho $k \times k$ (com $k = 4$), obtida da matriz Λ; e 1_k é o vetor coluna de dimensão igual ao número de estados operacionais, i.e., $k = 4$.

Através de (2.35) obtem-se os resultados da tabela 4.1, a qual apresenta o MTTF da configuração SA(2+1), e o respectivo λ^*_{SA}, em função de três valores do tempo médio para reparos (MTTR) e considerando $c = 0.96$. É importante notar que, em funçao das características particulares de funcionamento da estrutura SA(2+1), o MTTR praticamente

não exerce influência no MTTF e, consequentemente, no λ^{*}_{SA}. Na obtenção dos valores apresentados nesta tabela foi considerado $\lambda_{SA} = 6.3 \times 10^{-5} falhas/h$, conforme as taxas de falha reportadas em [HOL 91c].

MTTF(h)	λ^{*}_{SA}(falha/h)	MTTR(h)
395951.44	2.526×10^{-6}	0.5
395081.53	2.531×10^{-6}	1
393353.79	2.542×10^{-6}	2

Tabela 4.1 - MTTF e λ^{*}_{SA} em função do MTTR: valores referentes à configuração SA(2+1).

4.3 Análise da Cobertura

A análise apresentada nesta seção tem o propósito de ressaltar os efeitos da cobertura na indisponibilidade da estrutura MA/SA. Para tanto, é tomado como exemplo a indisponibilidade das funções de estabelecimento de chamada providas pela estrutura MA/SA (I_{esl}).

A indisponibilidade I_{esl} é obtida a partir das probabilidades estacionárias de ocupação dos estados referentes ao modelo da figura 4.2. Uma vez que I_{esl} representa a probabilidade de ocupação dos estados de falha, tem-se:

$$I_{esl} = P_{F_1} + P_{F_2} + P_{F_3} =$$

$$= [\frac{(\lambda_{MA} + \lambda_{SA})\overline{C}_A}{\mu'} + \frac{\mu''\lambda_{MA}\overline{C}_R}{\mu'(\lambda_{MA}\overline{C}_A + \mu'')} + \frac{\lambda_{MA}\overline{C}_R(\lambda_{MA}\overline{C}_A + \lambda_{SA})}{\mu(\lambda_{MA}\overline{C}_A + \mu'')} +$$

$$+ \frac{(\lambda_{MA}\overline{C}_A + \lambda_{SA})(2\lambda_{MA} + \lambda_{SA})}{\mu^2}]P_1 \quad , \tag{4.2}$$

onde

$$P_1 = [1 + \frac{2\lambda_{MA} + \lambda_{SA}}{\mu} + \frac{2\lambda^2_{MA}}{\mu^2} + \frac{3\lambda_{MA}\lambda_{SA}}{\mu^2} + \frac{\lambda^2_{SA}}{\mu^2} + \frac{(\mu + \lambda_{MA} + \lambda_{SA})\lambda_{MA}\overline{C}_R}{\mu(\mu'' + \lambda_{MA}\overline{C}_A)} +$$

$$+ \frac{(\lambda_{MA} + \lambda_{SA})\overline{C}_A}{\mu'} + + \frac{\mu''\lambda_{MA}\overline{C}_R}{\mu'(\lambda_{MA}\overline{C}_A + \mu'')}]^{-1} \quad .$$

A curva apresentada na figura 4.5 mostra a influência da cobertura na indisponibilidade da estrutura MA/SA. Nesta curva, $c_A = c_R$ e $\mu'' = 1/1440h$, o que equivale à existência do procedimento de comutação com um período equivalente a 60 dias. Além disso, esta curva corresponde à configuração usual (MA_a,MA_r,SA) e corresponde a $\mu' = 4$, $\mu = 1$, $\lambda_{MA} = 1.85 \times 10^{-4} falhas/h$, $\lambda_{MA} = 6.3 \times 10^{-5} falhas/h$.

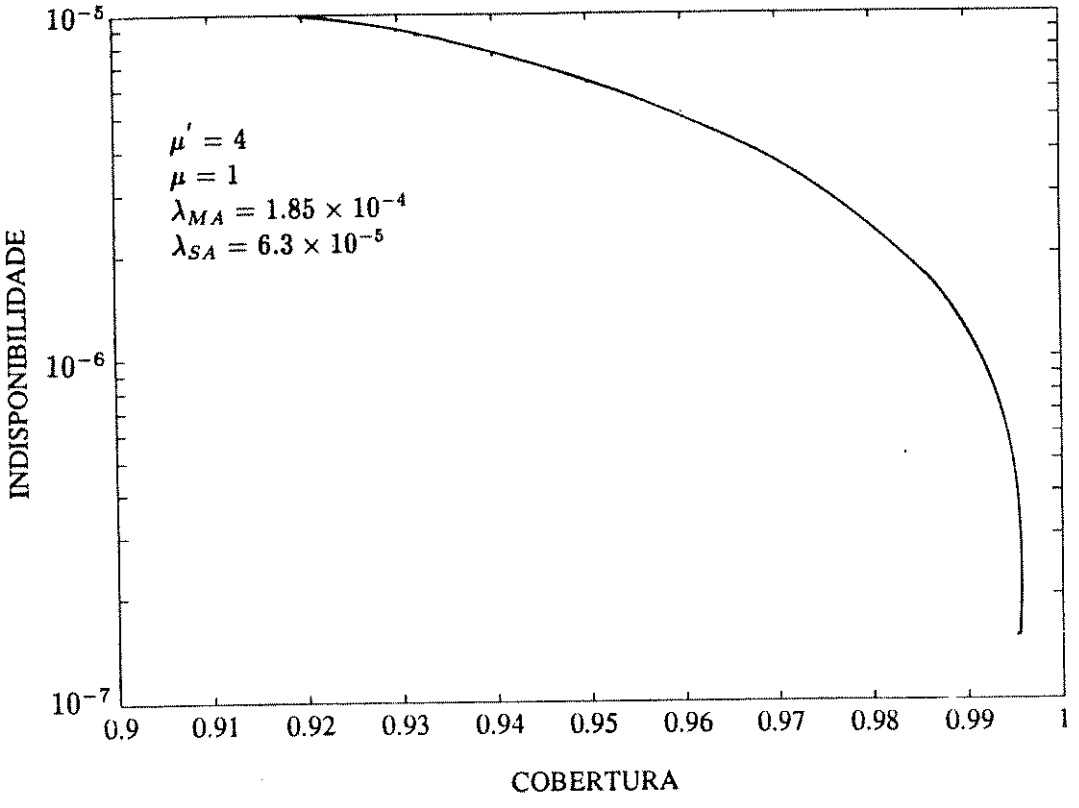


Figura 4.5 - I_{est} × cobertura ($c = c_A = c_R$).

A tabela 4.2 resume alguns valores de I_{est} em função de C_A e C_R, através dos quais também fica evidente o impacto da cobertura. Um exemplo deste impacto pode ser observado nos casos em que $C_A = C_R = 0.99$ e $C_A = C_R = 1$, onde a diferença entre os valores de I_{est} é aproximadamente 50 vezes. Nesta tabela, os valores de I_{est} foram obtidos com a configuração usual (MA_a,MA_r,SA) e com as mesmas taxas de falha utilizadas na obtenção das curvas.

Indisponibilidade (I_{est})	C_A	C_R
9.42×10^{-6}	0.9	0.95
6.88×10^{-6}	0.9	0.99
6.23×10^{-6}	0.9	1
9.34×10^{-6}	0.95	0.9
6.8×10^{-6}	0.99	0.9
6.16×10^{-6}	1	0.9
12.5×10^{-6}	0.9	0.9
5.03×10^{-6}	0.96	0.96
2.53×10^{-6}	0.98	0.98
1.28×10^{-6}	0.99	0.99
0.027×10^{-6}	1	1

Tabela 4.2 - I_{est}× cobertura: valores para $\mu' = 4$, $\mu = 1$, e $\mu'' = 1/1440h$.

4.4 Estimativas de Indisponibilidade

Esta seção apresenta estimativas de indisponibilidade hardware dos parâmetros avaliados neste estudo de caso, os quais correspondem a: Indisponibilidade das funções de estabelecimento de chamada e de tarifação, providas pela estrutura MA/SA; Indisponibilidade de um módulo LA; E indisponibilidade percebida por um usuário do nó. Os valores de indisponibilidade são apresentados em função de três valores do MTTR (0.5, 1 e 2 h), os quais refletem procedimentos possíveis de reparo, haja visto que o PSN é assistido localmente em termos de operação & manutenção.

Além disso, as estimativas de indisponibilidade são apresentads para os casos de existência e inexistência do procedimento de comutação periódica, quando a estrutura MA/SA é envolvida. Os valores para o caso de existência do procedimento correspondem a um período de comutação de 60 dias, ou seja, $\mu'' = 1/1440\,h$, o qual foi adotado nesta análise para ilustrar a flexibilidade do procedimento quanto à escolha do período de comutação, bem como resguardar o caráter conservativo da avaliação, uma vez que a indisponibilidade diminui para períodos de comutação menores.

Na obtenção dos valores apresentados a seguir, foram consideradas as taxas de falha reportadas em [HOL 91c], i.e., $\lambda_{MA} = 1.85 \times 10^{-4} falhas/h$ e $\lambda_{SA} = 6.3 \times 10^{-5} falhas/h$.

4.4.1 Indisponibilidade das funções de estabelecimento de chamada

A avaliação de indisponibilidade das funções de estabelecimento de chamada (I_{est}) é feita a partir de duas configurações da estrutura MA/SA: A configuração usual (MA_a,MA_r,SA) e a configuração (MA_a,MA_r,SA(2+1)).

A. Configuração usual (MA_a,MA_r,SA)

Para esta configuração, a indisponibilidade das funções de estabelecimento de chamada pode ser obtida conhecendo-se as probabilidades estacionárias de ocupação dos estados de falha ($P_{F_i}, i = 1,2,3$) do modelo da figura 4.2, uma vez que I_{est} é expressa pela equação (4.2).

A tabela 4.3 apresenta valores de I_{est} em função de três valores do MTTR, para os casos de inexistência e existência do procedimento de comutação periódica entre os MA's, $\mu'' = 0$ e $\mu'' = 1/1440\,h$, respectivamente.

Com relação aos resultados da tabela 4.3, é possível notar a influência da comutação periódica entre os MA's na indisponibilidade das funções de estabelecimento de chamada, o que se torna mais acentuado à medida que o MTTR aumenta. No caso do MTTR=2 h, por exemplo, I_{est} chega a ter uma redução de aproximadamente 12 vezes em seu valor, com a utilização da comutação periódica.

Indisponibilidade (I_{est})		MTTR (h)	I_{est1}/I_{est2}
I_{est1} ($\mu'' = 0$)	I_{est2} ($\mu'' = 1/1440h^{-1}$)		
18.8×10^{-6} (9.88 min./ano)	4.64×10^{-6} (2.44 min./ano)	0.5	4.05
36.4×10^{-6} (19.1 min./ano)	5.03×10^{-6} (2.64 min./ano)	1	7.24
71.6×10^{-6} (37.6 min./ano)	5.85×10^{-6} (3.08 min./ano)	2	12.24

Tabela 4.3 - $I_{est}\times$ MTTR : valores para a configuração (MA_a,MA_r,SA), $\mu' = 4$, $C_A = C_R = 0.96$ e $\lambda^*_{SA} = \lambda_{SA}$.

Para que seja possível verificar se este parâmetro atende aos requisitos de segurança de funcionamento, é necessário tecer algumas considerações preliminares. Como não existe uma diretriz a nível de recomendação que especifique um determinado limite máximo para a indisponibilidade das funções de estabelecimento de chamada de um PSN, é razoável comparar o fornecimento destas funções com as funções providas pela matriz de comutação de uma central telefônica. Neste sentido, a indisponibilidade das funções de estabelecimento de chamada de um PSN deve ser pelo menos igual à indisponibilidade especificada para uma central telefônica, uma vez que a comunicação de dados deve atender a padrões técnicos mais rígidos se comparados aos exigidos para a comunicação de sinais analógicos.

Segundo o CCITT [CCI 80], a indisponibilidade (I) de uma central telefônica não deve ser maior do que $2\,h$ em 40 anos de serviço (ou, em termos probabilísticos, $I < 5.8 \times 10^{-6}$). Desta forma, comparando os resultados da tabela 4.3 com este requisito, é possível observar que, para a configuração usual, ele só é atendido com a utilização do procedimento de comutação periódica.

B. Configuração (MA_a,MA_r,SA(2+1))

A indisponibilidade das funções de estabelecimento de chamada, referente a esta configuração, é obtida através do modelo da figura 4.2, considerando-se os valores de λ^*_{SA} apresentados na tabela 4.1.

A determinação de I_{est} é realizada de forma similar à adotada para a configuração usual, ou seja, através da expressão (4.2). Entretanto, neste caso é utilizado o λ^*_{SA} referente à estrutura SA(2+1). Como pode ser observado na tabela 4.4, os valores são apresentados em função do MTTR, e para cada valor do MTTR é utilizado o λ^*_{SA} correspondente.

Indisponibilidade (I_{est})		MTTR (h)	I_{est1}/I_{est2}
I_{est1} ($\mu'' = 0$)	I_{est2} ($\mu'' = 1/1440h$)		
3.42×10^{-6} (1.8 min./ano)	3.72×10^{-6} (1.96 min./ano)	0.5	0.92
5.9×10^{-6} (3.1 min./ano)	3.77×10^{-6} (1.98 min./ano)	1	1.56
10.9×10^{-6} (5.73 min./ano)	3.89×10^{-6} (2.04 min./ano)	2	2.8

Tabela 4.4 - $I_{est}\times$ MTTR : valores para a configuração (MA_a,MA_r,SA(2+1)), $\mu' = 4$ e $c_A = c_R = 0.96$.

A partir da tabela 4.4, é importante observar que:

- Para esta configuração, a existência da comutação periódica entre os MA's tem um impacto menor na indisponibilidade das funções de estabelecimento de chamada providas pela estrutura MA/SA. A explicação para a redução deste impacto reside no fato de que, nesta configuração, o $\lambda^*_{SA_i}$ é bem menor do que o referente à configuração usual. Desta forma, quando $\mu \rightarrow \mu^{'}$, os valores de indisponibilidade relativos aos casos de existência e inexistência da comutação periódica tendem a um mesmo patamar, chegando, inclusive, a uma situação onde a indisponibilidade para o caso em que há a comutação periódica é maior do que a indisponibilidade do caso em que não há.

  Este aspecto é contrário à intuição e pode ser melhor percebido a partir de uma inspeção mais detalhada do modelo da figura 4.2. Como pode ser observado, para $\mu^{''} = 0$, existem duas possibilidades de transição do estado 3 para os estados de falha (estados 7 e 8) com taxa de retorno μ mais lenta do que $\mu^{'}$ (associada ao estado 6). Esta condição "desfavorável" é compensada com a adoção do procedimento de comutação periódica, o qual possibilita a transição do estado 3 para o estado 6 que, apesar de ser um estado de falha, possui taxa de retorno $\mu^{'}$ mais rápida.

  Entretanto, esta compensação perde seu efeito quando μ se aproxima muito de $\mu^{'}$, uma vez que a taxa de retorno mais rápida deixa de existir (em termos comparativos) e $\mu^{''}$ passa a representar apenas uma possibilidade adicional de migrar do estado 3 para um estado de falha, cuja taxa de retorno está nivelada com as demais. Neste caso, a indisponibilidade do sistema pode até mesmo aumentar, dependendo da relação $\mu/\mu^{'}$. A partir dos resultados da tabela 4.4, é possivel notar que esta situação ocorre quando $\mu/\mu^{'}$=0.5, ou seja, o $MTTR\,(1/\mu)$ é da ordem de $0.5\,h$.

- Comparando estes resultados com o requisito de indisponibilidade para uma central telefônica, i.e., $I < 5.8 \times 10^{-6}$, a indisponibilidade das funções de estabelecimento de chamada, para esta configuração, atende tal requisito quando o MTTR é menor do que $1\,h$, ou quando o procedimento de comutação periódica é adotado.

4.4.2 Indisponibilidade das funções de tarifação

A avaliação da indisponibilidade das funções de tarifação (I_{tar}) de um PSN COMPAC é feita a partir da determinação das probabilidades estacionárias de ocupação dos estados de falha do modelo da figura 4.3, através de um procedimento similar ao apresentado na seção 4.3. A tabela 4.5 apresenta valores de I_{tar} em função do MTTR para os casos de existência e inexistência da comutação periódica dos MA's.

Indisponibilidade (I_{tar})		MTTR (h)	I_{tar1}/I_{tar2}
I_{tar1} ($\mu'' = 0$)	I_{tar2} ($\mu'' = 1/1440h$)		
47.2×10^{-6} (24.8 min./ano)	4.63×10^{-6} (2.43 min./ano)	0.5	10.2
93.4×10^{-6} (49.1 min./ano)	5.65×10^{-6} (2.97 min./ano)	1	16.5
186×10^{-6} (97.8 min./ano)	7.81×10^{-6} (4.1 min./ano)	2	23.8

Tabela 4.5 - $I_{tar}\times$ MTTR : valores para a configuração (MA_a,MA_r,SA), $\mu' = 4$ e $C_A = C_R = 0.96$.

Os resultados da tabela 4.5 ressaltam os efeitos significativos da comutação periódica na indisponibilidade das funções de tarifação providas pela estrutura. Para um MTTR=2 h, esta indisponibilidade chega a ser mais de 20 vezes menor que a indisponibilidade correspondente ao caso em que não há a comutação periódica ($\mu'' = 0$).

4.4.3 Indisponibilidade de um LA

Devido ao fato da arquitetura de um LA não apresentar redundância, sua indisponibilidade (I_{LA}) é obtida a partir de um modelo com dois estados: um, operacional e outro, de falha, resultando na equação mostrada a seguir.

$$I_{LA} = \frac{MTTR}{MTBF + MTTR}$$

onde λ é a taxa de falha do módulo LA.

Esta análise considera a indisponibilidade de um usuário conectado a um LA. A conexão deste usuário é realizada através de uma LB, e o insucesso desta conexão pode ocorrer com falha desta e/ou de uma das demais unidades hardware do LA. Desta forma, objetivando atribuir um caráter conservativo à análise, foi considerada a situação de um usuário conectado a uma LB16, já que, dentre os três tipos de LB, esta é a que apresenta maior taxa de falha.

A tabela 4.6 apresenta valores de indisponibilidade hardware de um LA, sob o ponto de vista de um usuário, em função do MTTR.

Indisponibilidade (I_{LA})	MTTR (h)
3.55×10^{-5} (18.66 min./ano)	0.5
7.10×10^{-5} (37.32 min./ano)	1
14.2×10^{-5} (74.64 min./ano)	2

Tabela 4.6 - Indisponibilidade de um LA em função do MTTR.

4.4.4 Indisponibilidade percebida por um usuário

Um determinado usuário de um PSN COMPAC pode perceber a indisponibilidade de um enlace e a indisponibilidade das funções de estabelecimento de chamada. A avaliação associada à percepção destes dois aspectos é apresentada a seguir.

Nesta avaliação, um enlace entre dois usuários envolve os dois LA's situados nas pontas do caminho de comunicação, a estrutura MA/SA e, naturalmente, o meio de transmissão. A análise aqui apresentada não inclui a indisponibilidade do meio de transmissão, uma vez que muitos fatores externos ao produto sob avaliação, e de difícil comensurabilidade, precisariam ser considerados.

A. Indisponibilidade de um enlace

Entre os serviços oferecidos pelo sistema COMPAC, existe a possibilidade de formação de um Circuito Virtual Permanente (CVP), onde dois usuários podem estabelecer uma chamada permanente. Neste caso, a indisponibilidade do hardware envolvido num enlace entre os dois usuários de um CVP, representada por I_{enl}, constitui um parâmetro importante, que, por sua vez, é determinado pela expressão (4.3).

$$I_{enl} = 1 - (1 - I_{LA_a})(1 - I_{LA_b}) \tag{4.3}$$

onde I_{LA_a} e I_{LA_b} correspondem, respectivamente, à indisponibilidade dos LA's aos quais os assinantes a e b estão conectados. A expressão (4.3) não inclui a indisponibilidade da estrutura MA/SA pelo fato de que, uma vez estabelecido o CVP, esta estrutura não exerce mais influência na manutenção do caminho de comunicação entre os usuários envolvidos. A tabela 4.7 apresenta valores de I_{enl} em função do MTTR.

Indisponibilidade (I_{enl})	MTTR (h)
7.1×10^{-5} (37.32 min./ano)	0.5
14.2×10^{-5} (74.64 min./ano)	1
28.4×10^{-5} (149.3 min./ano)	2

Tabela 4.7 - Indisponibilidade de um enlace em função do MTTR.

B. Indisponibilidade das funções de estabelecimento de chamada percebida por um usuário

Um determinado usuário fica impossibilitado de estabelecer qualquer chamada quando o LA ao qual está conectado e/ou as funções de estabelecimento de chamada providas pela estrutura MA/SA não se encontram disponíveis. Desta forma, esta indisponibilidade, representada por I^*_{est}, é dada pela relação apresentada em (4.4).

$$I^*_{est} = 1 - (1 - I_{LA})(1 - I_{est}) \tag{4.4}$$

onde I_{LA} é a indisponibilidade de um LA e I_{est} é a indisponibilidade das funções de estabelecimento de chamada providas pela estrutura MA/SA, apresentada na subseção 4.4.1.

A tabela 4.8 apresenta valores de I^*_{est} correspondentes aos valores de I_{est} para a configuração (MA_a,MA_r,SA), mostrados na tabela 4.3, e aos valores de I_{LA} apresentados na tabela 4.6.

Indisponibilidade (I^*_{est})		MTTR
$\mu'' = 0$	$\mu'' = 1/1440h$	(h)
5.43×10^{-5} (28.5 min./ano)	4.01×10^{-5} (21.1 min./ano)	0.5
10.74×10^{-5} (56.4 min./ano)	7.6×10^{-5} (39.9 min./ano)	1
21.36×10^{-5} (112.3 min./ano)	14.78×10^{-5} (77.7 min./ano)	2

Tabela 4.8 - $I^*_{est}\times$ MTTR: valores para a configuração (MA_a,MA_r,SA)

Através da tabela 4.8, é possível notar que:

- A indisponibilidade é menor para o caso em que a estrutura MA/SA tem a comutação periódica entre os MA's (valores apresentados na coluna em que $\mu'' = 1/1440h$).
- A indisponibilidade do LA exerce uma influência bem maior no cômputo da indisponibilidade percebida pelo usuário, uma vez que os valores desta tabela estão mais próximos dos valores apresentados na tabela 4.6, do que dos valores da tabela 4.3. Isto se deve ao fato de que o módulo LA não possui redundância, portanto apresenta uma indisponibilidade bem maior do que a das funções de estabelecimento de chamada providas pela estrutura MA/SA.

Para verificação do atendimento aos requisitos de segurança de funcionamento, é necessário adotar um procedimento comparativo, similar ao descrito na subseção 4.4.1, sendo que, no presente caso, balizando-se em dois requisitos: *i*) a indisponibilidade de terminal de uma central telefônica, i.e., uma indisponibilidade menor do que 10^{-4} [CCI 80]; e *ii*) a indisponibilidade de uma terminação de assinante ligado a uma central RDSI, i.e., menor ou igual a $30\,min.$ ao ano ($I \leq 5.7 \times 10^{-5}$) [CCI 88b].

Assim, comparando os resultados da tabela 4.8 com a indisponibilidade de terminal recomendada para uma central telefônica, nota-se que este requisito é atendido para um MTTR menor do que $1\,h$, independentemente da utilização do procedimento de comutação periódica. Por outro lado, comparando os resultados obtidos com a indisponibilidade recomendada para uma terminação de central RDSI, a qual constitui um requisito muito mais condizente com o tipo de serviço fornecido por um PSN, é possível observar que este requisito é atendido para um MTTR de até $30\,min.$, ou para um MTTR próximo a $1\,h$, quando o procedimento de comutação periódica é usado.

É importante ressaltar que, para a indisponibilidade das funções de estabelecimento de chamada percebida por um usuário, a utilização do procedimento de comutação tem um impacto menor se comparado aos demais parâmetros. Tal aspecto é explicado pelo fato de

que esta indisponibilidade é fortemente condicionada, conforme já mencionado, pela indisponibilidade do LA, o qual não apresenta redundância, e, consequentemente, independe do processo de comutação periódica.

4.5 Conclusões

Este capítulo apresentou o processo de avaliação da segurança de funcionamento de um PSN COMPAC, ilustrando a aplicação de algumas das técnicas descritas no capítulo 2, e destacando o impacto que o procedimento de comutação periódica exerce sobre a indisponibilidade dos parâmetros avaliados. Baseado no que foi apresentado, é possível tecer os seguintes comentários:

- Esta avaliação possibilitou, além da confrontação das estimativas de indisponibilidade obtidas com os requisitos de segurança de funcionamento, detectar como aspectos ligados a particularidades operacionais do PSN, como a cobertura e o procedimento de comutação periódica, podem influenciar os parâmetros avaliados. Apoiado nesta constatação, foi possível apresentar sugestões que conduzem a um aumento da qualidade de serviço do sistema.

- A adoção de um procedimento formal de comutação periódica entre os MA's tem um impacto significativo na indisponibilidade das principais funções providas pela estrutura MA/SA. Um aspecto importante deste impacto está relacionado com o fato de que ele apresenta flexibilidade quanto à escolha do período da comutação, podendo, inclusive, ser da ordem de dezenas de dias. Desta forma, além de reduzir consideravelmente a indisponibilidade destas funções, principalmente no que diz respeito à tarifação, o procedimento de comutação é de fácil implantação, não requer implementações ou alterações no sistema, e praticamente não implica em custos adicionais.

- A influência do tempo médio para reparos (MTTR) na indisponibilidade do hardware envolvido no desempenho das funções do PSN merece atenção especial. A escolha de uma estratégia de manutenção que conduza a um MTTR de curta duração implica, diretamente, na obtenção de uma indisponibilidade menor.

- A utilização de um fator de cobertura de 96% corresponde ao propósito de resguardar os resultados obtidos, a partir de uma avaliação baseada em estimativas conservativas. Conforme apresentado, a título ilustrativo, na seção 4.3, a cobertura tem grande influência na indisponibilidade da estrutura MA/SA. Consequentemente, o levantamento do fator de cobertura associado a esta estrutura, conforme metodologia descrita no capítulo 3, permitiria avaliações com resultados mais precisos. Além disso, a determinação do fator de cobertura possibilitaria, também, a avaliação do desempenho de funcionamento dos mecanismos de supervisão.

Conclusão Geral

O capítulo 1 introduziu os conceitos básicos da segurança de funcionamento, destacando a avaliação de seus atributos como forma de garantir o desenvolvimento de sistemas em consonância com os requisitos de qualidade de serviço. O capítulo 2 apresentou uma ampla gama de técnicas e ferramentas utilizadas atualmente na avaliação dos atributos de segurança de funcionamento. Os capítulos 3 e 4 mostraram o processo de avaliação da segurança de funcionamento para dois sistemas de telecomunicações, ilustrando a aplicação de algumas das técnicas descritas no capítulo 2. Além disso, o capítulo 3 descreveu uma proposta de metodologia para levantamento do fator de cobertura, e o capítulo 4 destacou como um determinado procedimento operacional pode contribuir significativamente para elevar o nível de segurança de funcionamento do sistema.

A partir do que foi apresentado, é possível constatar que a avaliação de segurança de funcionamento permite não só verificar se os sistemas estão sendo desenvolvidos de acordo com os requisitos de qualidade de serviço, como também detectar pontos a partir dos quais pode-se sugerir meios que elevem o nível de segurança de funcionamento de tais sistemas.

A avaliação apresentada no capítulo 4 ilustra muito bem como os resultados provenientes da avaliação de um parâmetro de segurança de funcionamento podem contribuir para o aumento da qualidade dos serviços prestados pelo sistema, tanto em termos da qualidade percebida pelo usuário, quanto no que se refere à qualidade das facilidades de operação e controle empregadas pela empresa administradora do serviço. Conforme evidenciam os resultados apresentados nesta avaliação, a simples adoção de um procedimento de comutação periódica entre as unidades ativa e reserva da estrutura de controle do nó pode aumentar significativamente a disponibilidade das funções providas por esta estrutura. Os resultados mostram ainda que as melhorias advindas com este procedimento são mantidas mesmo para períodos de comutação da ordem de vários dias, o que aumenta a facilidade de implantação do procedimento.

A avaliação apresentada no capítulo 3 mostra outro aspecto digno de nota, o qual está relacionado com a utilização de técnicas de avaliação que melhor se adaptam às necessidades de projeto. Como pôde ser observado neste capítulo , a elaboração de uma metodologia para levantamento do fator de cobertura possibilitou conciliar rapidez de resposta com utilização dos recursos disponíveis ao longo do desenvolvimento do produto. Além disso, o emprego desta metodologia leva à obtenção de resultados conservativos, cuja precisão é satisfatória em termos de sistemas de telecomunicações voltados a aplicações comerciais. Desta forma, é possível dispor de um método prático e eficiente, que condiciona a obtenção de estimativas mais realistas dos parâmetros de segurança de funcionamento.

A utilização desta metodologia para levantamento do fator de cobertura permite, ainda, aferir a atuação dos mecanismos de supervisão, os quais representam elementos de suma importância na implantação de métodos eficazes de tolerância a defeitos. Com isto é possível descobrir mecanismos que não estejam funcionando conforme o esperado, e que,

por este motivo, requeiram aprimoramentos no sentido de garantir o nível de segurança de funcionamento almejado.

Através da avaliação de segurança de funcionamento, é possível obter também informações valiosas à escolha de uma estratégia de reparos condizente com o nível de segurança de funcionamento pretendido. Com os resultados da avaliação percebe-se, quantitativamente, a influência que o tempo envolvido num reparo exerce sobre os parâmetros de segurança de funcionamento. Apesar deste aspecto ser intuitivamente esperado, a determinação adequada da relação *tempo de reparo* X *segurança de funcionamento* só é alcançável com o conhecimento quantitativo desta influência. Tal quantificação, por sua vez, é proporcionada pela avaliação de segurança de funcionamento, como pode ser observado nos capítulos 3 e 4.

De uma forma geral, todos estes aspectos observados se comportam como agentes catalizadores do processo de garantia de segurança de funcionamento, cujo objetivo é proporcionar o desenvolvimento de sistemas capazes de fornecer serviços com uma qualidade consonante com o tipo de aplicação a que se destinam.

Com relação à continuação e aprofundamento deste trabalho, podemos destacar as seguintes perspectivas:

- Investigar técnicas de elaboração de modelos com grande número de estados (da ordem de 10^4 estados), e.g., Redes de Petri Estocásticas e Inteligência Artificial, que atendam a uma relação custo × desempenho satisfatória. Com estas técnicas, torna-se factível a modelagem de sistemas complexos, que requeiram avaliações com alto grau de precisão.

- Estender a metodologia de determinação do fator de cobertura para os defeitos transientes, de forma a fornecer uma análise mais abrangente da cobertura, conciliando os mesmos fatores de praticidade e aproveitamento dos recursos existentes, contemplados para o caso dos defeitos permanentes. O interessante é que esta nova metodologia também conduza a resultados com precisão satisfatória, em se tratando de sistemas comerciais de telecomunicações, e que não requeira o desenvolvimento de subprodutos, como, por exemplo, os exigidos pelos procedimentos de injeção de defeitos.

- Contemplar os atributos de segurança de funcionamento não incluídos nesta abordagem, como os relacionados com *performability* e com a segurança de funcionamento contra riscos (*safety*) e intrusões (*security*), no sentido de proporcionar a avaliação para outros tipos de sistema.

O acréscimo de elementos como estes contribui no sentido de fomentar avaliações de segurança de funcionamento mais completas e a custos reduzidos, além de estendê-las para sistemas não pertencentes ao domínio das telecomunicações. Desta maneira, é possível trilhar uma linha evolutiva que vá ao encontro da tendência de expansão do conceito de segurança de funcionamento.

Esta expansão conceitual é justificada, sobretudo, pela sofisticação crescente dos sistemas computacionais, a qual induz a necessidade de considerar um número cada vez maior

de tipos de entraves à segurança de funcionamento, tais como: falhas de concepção software, falhas de concepção de sistema, falhas de procedimento operacional, falhas intencionais, etc. Como decorrência a esta diversificação da natureza dos entraves, surge, naturalmente, duas necessidades: dar um novo enfoque à avaliação dos atributos de segurança de funcionamento, e ampliar a gama destes atributos, no sentido de adequar as avaliações de segurança de funcionamento às exigências emergentes.

Bibliografia

[**ARL 89a**] J.Arlat et al., "SURF-2: Specification de Conception", Rapport LAAS No.89-098, Mars 1989.

[**ARL 89b**] J.Arlat, Y.Crouzet, J.C.Laprie, "Fault Injection for Dependability Validation of Fault-Tolerant Computing Systems". *Proc. 19th IEEE Int. Symp. on Fault Tolerant Computing (FTCS-19)*, Chicago, June 1989, pp.348-355.

[**ARN 73**] T.F.Arnold, "The Concept of Coverage and its Effect on the Reliability Model of a Repairable System", *IEEE Trans. on Computer*, vol. C-22, Mar. 1973, pp.251-254.

[**ARR 87**] F.Arruda Jr. et al.," O Sistema COMPAC para Redes de Comunicação de Dados", Anais do 5º Simpósio Brasileiro de Redes de Computadores, São Paulo - SP, p.146-161, Abril 1987.

[**BAR 65**] R.E.Barlow & F.Proscham. *Mathematic Theory of Reliability.* New York, Wiley, 1965.

[**BAZ 86**] I.Bazovsky Sr. & I.Bazovsky Jr., "Fault Coverage of Intelligent Switching Networks", *IEEE Journal on Selected Areas in Communications*, Vol.SAC-4, No.6, October 1986, pp.1138-1142.

[**BOB 86**] A.Bobbio & K.S.Trivedi, "An Aggregation Technique for the Transient Analysis of Stiff Markov Chains", *IEEE Trans. on Computers*, Vol.C-35, No. 9, September 1986, pp. 803-814.

[**BUZ 70a**] J.A.Buzacott, "Network Approaches to Finding the Reliability of Repairable Systems", *IEEE Trans. on Reliability*, Vol.R-19, No.4, November 1970, pp.140-146.

[**BUZ 70b**] J.A.Buzacott, "Markov Approach to Finding Failure Times of Repairable Systems", *IEEE Trans. on Reliability*, Vol.R-19, No.4, Nov. 1970, pp.128-133.

[**CCI 80**] CCITT - International Telegraph and Telephone Consultative Committee. GAS 6, No. 51-E , Mar. 1980.

[**CCI 84**] ITU-CCITT, *Handbook of Quality of Service.* 1984

[**CCI 86**] ITU-CCIR, *Recommendations and Reports of the CCIR. Vol IV part I: Fixed Satellite Service.* Geneva, ITU, 1986.

[**CCI 88a**] ITU-CCITT, *Telephone Network and ISDN: Quality of Service, Network Management and Traffic Engineering.* Recommendations E.401-E.880, vol.II-Fascicle II.3, IX Plenary Assembly, Melbourne, Nov. 1988.

[**CCI 88b**] ITU-CCITT, "Rec. Q.543:Digital Exchange Performance Design Objectives". Blue Book, Vol.VI, Fas. VI.5, 1988.

[ÇIN 75] E.Çinlar, *Introduction to Stochastic Process.* Prentice-Hall, Englewood Cliffs, N.J., 1975.

[DUG 89] J.B.Dugan & K.S.Trivedi, "Coverage Modeling for Dependability Analysis of Fault-Tolerant Systems", *IEEE Trans. on Computers*, Vol. 38, No.6, June 1989, pp. 775-787.

[FEI 86] I.A.Feigenbaum, "Reliability of Commercial Communication Satellite Systems, *IEEE Journal on Selected Areas in Communication*, Vol.SAC-4, No.7, October 1986, pp.1034-1038.

[FRA 89] J.F.M.S.Franco, R.A.Gonçalves, E.S.Libardi, G.M.de Holanda, A. Barbieri, "Aquisição e Manutenção de Sincronismo no SAMSAT", *Anais do 7º Simpósio Brasileiro de Telecomunicações*, Florianópolis-SC, p.102-108 , Setembro 1989.

[GEI 83] R.Geist & K.S.Trivedi, "Ultra-High Reliability Prediction for Fault-Tolerant Computer Systems", *IEEE Trans. on Computers*, vol. C-32, No.12, December 1983, pp. 1118-1127.

[GEI 90] R.Geist & K.S.Trivedi, "Reliability Estimation of Fault-Tolerant Systems: Tools and Techniques", *Computer*, vol. 23, No. 7, July 1990, pp. 52-61.

[GON 86] R.A.Gonçalves, "Sistema de Comunicação de Dados via Satélite Utilizando a Técnica AMDT Multifrequência", *Anais do 4º Simpósio Brasileiro de Telecomunicações*, Rio de Janeiro - RJ, p.173-177 , Setembro 1986.

[GON 87] R.A.Gonçalves, J.F.M.S.Franco, E.S.Libardi, "Filosofia e Mecanismos de Tolerância a Falhas Utilizados no SAMSAT", *Anais do II Simpósio de Computadores Tolerantes a Falhas*, Campinas-SP, p.63-74, Agosto 1987.

[GOY 86] A.Goyal et al., "The System Availability Estimator", *Proc. 16th Int. Symp. Fault-Tolerant Computing*, CS Press, Los Alamitos, Calif., July 1986, pp.84-89.

[GRO 84] D.Gross & D.R.Miller, "The Randomization Technique as a Modeling Tool and Solution Procedure for Transient Markov Processes", *Operations Research*, Vol.32, No.2, March-April 1984, pp.343-361.

[GRU 86] J.G.Gruber, E.Abdou, P.Richards and G.Williams, "Quality-of-Service in Evolving Telecommunications Networks", *IEEE Journal on Selected Areas in Communications*, Vol.SAC-4, No. 7, October 1986, pp.1084-1089.

[HOL 88] G.M.de Holanda, J.F.M.S.Franco, R.A.Gonçalves, "Disponibilidade do SAMSAT - Sistema de Comunicação de Dados via Satélite Usando a Técnica AMDT", *Anais do 6º Simpósio Brasileiro de Telecomunicações*, Campina Grande-PB, p.09-13, Setembro 1988.

[HOL 89] G.M. de Holanda & J.F.M.S.Franco, "Análise da Cobertura na Estrutura de Referência e Controle do SAMSAT", *Anais do III Simpósio de Computadores Tolerantes a Falhas*, Rio de Janeiro-RJ, p.197-213 , Setembro 1989.

[**HOL 91a**] G.M.de Holanda & J.F.M.S.Franco,"Reliability Aspects in Satellite Data Communication Equipment: A Case Study", *Proc. Reliability'91 Conference*, London-UK, June 1991, pp.299-313.

[**HOL 91b**] G.M. de Holanda, E.Moschim, J.Moreira de Souza, "Avaliação do Impacto de Supervisões Periódicas na Indisponibilidade de um Sistema Redundante". *Anais do IV Simpósio de Computadores Tolerantes a Falhas*, Gramado-RS, p.255-266, Outubro 1991.

[**HOL 91c**] G.M.de Holanda, A.V.Vinhas, J. Moreira de Souza, "Avaliação de Confiabilidade do COMPAC". Relatório Técnico, CPqD-TELEBRÁS, Campinas-SP, 1991.

[**HOL 92**] G.M.de Holanda & A.V.Vinhas, "Reliability Evaluation: Application on a Packet Switching Node". To be published in *Proc. European Safety and Reliability Conference*, Kopenhagen-DK, June 1992.

[**IRL 88**] E.A.Irland, "Assuring Quality and Reliability of Complex Electronic Systems: Hardware and Software", *Proc. of the IEEE* (invited paper), Vol.76, No.1, January 1988, pp.5-18.

[**JEN 53**] A.Jensen, *Markoff Chains as an Aid in the Study of Markoff Processes.* Skand Aktuarietidskr. 36, 1953.

[**KLE 75**] L.Kleinrock, *Queueing Systems: Vol. I.* John Wiley & Sons, New York, 1975.

[**KUM 80**] A.Kumar & M.Agarwal, "A Review of Standby Redundant Systems". *IEEE Trans. on Reliability*, vol. R-29, No. 4, October 1980, pp. 290-294.

[**LAN 78**] C.Landrault & J.C.Laprie, "SURF - A Program for Modeling and Reliability Prediction for Fault-Tolerant Computing Systems". *Information Technology*, J.Moneta Ed. Amsterdam, The Netherlands: North-Holland, 1978.

[**LAP 76**] J.C.Laprie, " On Reliability Prediction of Repairable Redundant Digital Structures." *IEEE Trans. on Reliability*, Vol.R-25, No. 4, October 1976, pp.256-258.

[**LAP 84**] J.C.Laprie, "Dependability Modeling and Evaluation of Software-and-Hardware Systems", *2nd GI/NTG/GMR Conference on Fault-Tolerant Computing* (invited paper), Bonn, September 1984.

[**LAP 86**] J.C.Laprie, "The Dependability Approach to Critical Systems". *Proc. 5th International Workshop on Trends in Safe Real Time Computer Systems (SAFECOMP'86)*, Sarlat, France, Oct. 14-17, 1986.

[**LAP 90**] J.C.Laprie, "Dependability: Basic Concepts and Associated Terminology". Report LAAS No. 90.055, Mar. 1990.

[**LEI 87**] J.G.B.Leite & O.G.Loques Filho, " Introdução à Tolerância a Falhas, Cap.4: Software", *Mini Curso do II Simpósio de Computadores Tolerantes a Falhas*, Campinas-SP, p.99-171, Agosto 1987.

[**MAR 91**] M.R.Bastos Martini, "Qualidade de Serviço de Sistemas Computacionais: Avaliação de Segurança de Funcionamento Quanto a Falhas de Concepção Hardware e Software". Tese de Doutorado, FEE-Unicamp, 1991.

[**MCG 85**] J.McGough, M.Smotherman, K.S.Trivedi, "The Conservatiness of Reliability Estimates Based on Instantaneous Coverage", *IEEE Trans. on Computers*, vol. C-34, No.7, July 1985, pp. 602-609.

[**MEY 78**] J.F.Meyer, "On Evaluating the Performability of Degradable Computing Systems", *Proc. 8th IEEE Int. Symp. on Fault-Tolerant Computing (FTCS-8)*, Toulose, France, June 1978, pp.44-49.

[**MIL 86**] MIL-HDBK-217 E *Military Standardization Handbook: Reliability Prediction of Electronic Equipment.* Department of Defense of the United States of America, 1986.

[**MIL 91**] MIL-HDBK-217 F *Military Standardization Handbook: Reliability Prediction of Electronic Equipment.* Department of Defense of the United States of America, 1991.

[**MOL 78**] C.Moler & C.Van Loan, "Nineteen Dubious Ways to Compute the Exponential of a Matrix", *SIAM REVIEW*, Vol.20, No.4, October 1978, pp.801-836.

[**NEL 90**] V.P.Nelson, "Fault-Tolerant Computing: Fundamental Concepts". *Computer*, Vol. 23, No.7, July 1990, pp. 19-25.

[**ODA 81**] Y.Oda et al., "Reliability and Performance Evaluation of Self-Reconfigurable Systems with Periodic Maintenance", *Proc. FTCS-11*, June 1981, pp.142-147.

[**OSA 76**] S.Osaki & T.Nakagawa, "Bibliography for Reliability and Availability of Stochastic Systems", *IEEE Trans. on Reliability*, Vol. R-25, October 1976, pp. 284-287.

[**PIT 89**] J.M.Pitsch, "SAMSAT - A Medium Data-Rate TDMA System". Proc. VIII ICDSC, French West Indies, Section A9, 1989.

[**RAC 79**] Reliability Analysis Center, *Reliability Design Handbook.* IIT Research Institute, 1979.

[**REI 88**] A.Reibman & K.Trivedi, "Numerical Transient Analysis of Markov Models", *Comput. Opns. Res.*, Vol.15, No.1, 1988, pp.19-36.

[**REI 91**] A.L.Reibman & M.Veeraraghavan, "Reliability Modeling: An Overview for System Designers". *Computer*, Vol. 24, No. 4, April 1991, pp. 49-57.

[**RIC 88**] J.S.Richters & C.A.Dvorak, "A Framework for Defining the Quality of Communications Services", *IEEE Communications Magazine*, October 1988, pp.17-23.

[**SAH 87**] R.Sahner & K.S.Trivedi, "Reliability Modeling Using SHARPE", *IEEE Trans. on Reliability*, Vol.R-36, No.2, June 1987, pp.186-193.

[**SHO 68**] M.L.Shooman, *Probabilistic Reliability: An Engineering Approach.* McGraw-Hill, 1968.

[**SIE 84**] D.P.Siewiorek, "Architecture of Fault-Tolerant Computers", *Computer*, Vol. , No. , August 1984, pp.9-18.

[**SIL 89**] E.S.e Silva & H.R.Gail, "Calculating Availability and Performability Measures of Repairable Computer Systems Using Randomization", *Journal of the Association for Computing Machinery*, Vol.36, No.1, January 1989, pp.171-193.

[**SIL 90**] E.S.e Silva & H.R.Gail, "Analyzing Scheduled Maintenance Policies for Repairable Computer Systems", *IEEE Trans. on Computers*, Vol.39, No.11, November 1990, pp.1309-1324.

[**SMI 88**] R.M.Smith, K.S.Trivedi, A.V.Ramesh, "Performability Analysis: Measures, an Algorithm, and a Case Study", *IEEE Trans. on Computers*, Vol.37, No.4, April 1988, pp.406-417.

[**SOU 87**] J.Moreira de Souza & M.R.B.Martini, "Introdução à Tolerância a Falhas, Cap.1: Avaliação da Confiabilidade de Sistemas", *Mini Curso do II Simpósio de Computadores Tolerantes a Falhas*, Campinas-SP, pp.5-62, Agosto 1987.

[**SOU 88**] J.Moreira de Souza, "Techinique d'Agregation dans un PdM", Rapport d'etude SURF-2, LAAS, Janvier 1988.

[**STA 87**] G.E.Stark, "Dependability Evaluation of Integrated Hardware/Software Systems", *IEEE Trans. on Reliability*, Vol.R-36, No.4, October 1987, pp.440-444.

[**STR 88**] K.Strandberg, "Field Dependability Evaluation Principles", *IEEE Journal on Selected Areas in Communications*, Vol.6, No.8, October 1988, pp.1330-1337.

[**TRI 82**] K.S.Trivedi, *Probability and Statistics with Reliability, Queuing, and Computer Science Applications.* Prentice-Hall, Englewood Cliffs, N.J.,1982.

[**TRI 84**] K.S.Trivedi & J.B.Dugan, "Modeling Imperfect Coverage in Fault-Tolerant Systems", *Proc. 14th Int. Conf. on Fault Tolerant Computers*, 1984, pp. 77-82.

[**WNG 77**] Y.W.Ng & A.Avizienis, "ARIES-An Automated Reliability Estimation System", *Proc. 1977 Annu. Reliability, Maintenability Symp.*, January 1977.

[**YAK 86**] Y.W.Yak et al., "The Effect of Incomplete and Deleterious Periodic Maintenance on Fault-Tolerant Computer Systems", *IEEE Trans on Reliability*, Vol.R-35, No.1, April 1986, pp.85-90.

Apêndice A

Modelo Semi-Markoviano para Avaliação da Indisponibilidade de um Sistema *Hot-Standby* 1:1

A análise apresentada neste apêndice tem como objetivo levantar o nível de influência que a hipótese de distribuição exponencial do tempo de supervisão exerce sobre a avaliação de indisponibilidade de um sistema com redundância *hot standby*. Esta análise se deve ao fato de que tal distribuição pode não ser, necessariamente, exponencial.

Conforme apresentado em [LAP 76], a hipótese de distribuição exponencial é válida quando a relação entre as taxas de reparo e de falha é alta (entre 10^3 e 10^6). Entretanto, em termos práticos, é interessante considerar também a possibilidade de períodos de supervisão extremamente altos, onde a relação μ/λ pode assumir valores muito menores do que 10^3. Face a esta necessidade, cabe verificar se a hipótese de distribuição exponencial, a qual simplifica o processo de modelagem, também é aceitável para $\mu/\lambda << 10^3$.

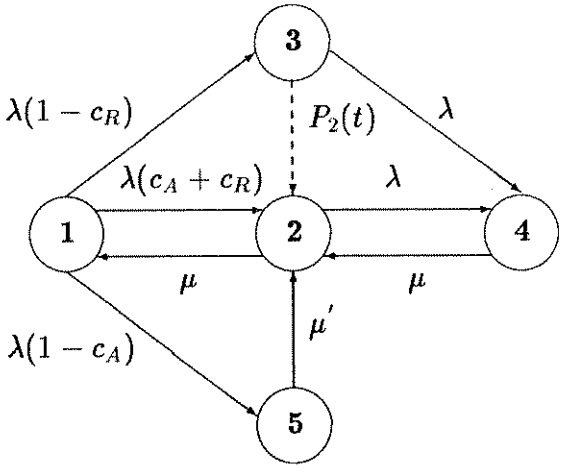


Figura A.1 - Modelo semi-markoviano para um sistema redundante (*hot standby 1:1*) com supervisões periódicas.

A figura A.1 corresponde a um modelo que descreve o comportamento do sistema *hot standby* 1:1 com supervisões periódicas, apresentado na subseção 2.2.4, segundo um processo semi-markoviano com tempo contínuo [KLE 75]. Este modelo apresenta os mesmos

estados e a mesma notação descritos para o modelo da figura 2.5, com exceção da taxa de transição $P_2(t)$ entre os estados 3 e 2, que neste caso está associada a um tempo de transição com distribuição geral, contabilizado a partir do instante em que o estado 3 é alcançado.

Como primeiro passo de um procedimento que permite o acesso à indisponibilidade deste sistema, determinemos as probabilidades que governam as transições do modelo de Markov embutido (*imbedded*) [HOW 71], associado ao modelo da figura A.1. Desta forma, as probabilidades de transição entre os estados i e j, (a^*_{ij}), do modelo embutido são dadas por:

$$a^*_{12} = \frac{c_A+c_R}{2} \quad , \quad a^*_{13} = \frac{1-c_R}{2} \quad , \quad a^*_{15} = \frac{1-c_A}{2}$$

$$a^*_{21} = \frac{\mu}{\lambda+\mu} \quad , \quad a^*_{24} = \frac{\lambda}{\lambda+\mu} \quad , \quad a^*_{32} = P_2$$

$$a^*_{34} = 1 - P_2 \quad , \quad a^*_{42} = 1 \quad , \quad a^*_{52} = 1$$

Cálculo de P_2:

Seja X_{ij} a variável aleatória (v.a.) que representa o tempo de transição do estado i para o estado j do modelo da figura A.1. Consequentemente, P_2 pode ser obtida da seguinte forma:

$$P_2 = \int_0^\infty P[t < X_{32} \leq t + \Delta t].P[X_{34} > t]dt = \int_0^\infty f_{32}(t)(1 - F_{34}(t))dt \quad , \qquad (A.1)$$

onde $f_{32}(t)$ é a função densidade da v.a. X_{32} e $F_{34}(t)$ é a função distribuição da v.a. X_{34}. Como X_{34} é exponencialmente distribuída com taxa λ, tem-se:

$$P_2 = \int_0^\infty f_{32}(t)e^{-\lambda t}dt \quad . \qquad (A.2)$$

Multiplicando ambos os lados da equação (A.2) por e^{-st}, obtem-se:

$$P_2 e^{-st} = \int_0^\infty e^{-st}e^{-\lambda t}f_{32}(t)dt = \int_0^\infty e^{-(s+\lambda)t}f_{32}(t)dt = f^*_{32}(s)|_{s=s+\lambda} \quad , \qquad (A.3)$$

onde $f^*_{32}(s)|_{s=s+\lambda}$ é a transformada de Laplace de $f_{32}(t)$ em $s = s + \lambda$.

Assumindo um procedimento de supervisão com período constante e igual a T, o qual corresponde ao procedimento mais usual do ponto de vista de sua implantação, e utilizando a propriedade de que

$$f_1(t) * f_{32}(t) = \delta(t - T) \quad , \qquad (A.4)$$

onde

$f_1(t) \rightarrow$ é a função densidade da v.a. X que representa, no modelo da figura A.1, a condicão do tempo de transição do estado 1 para o estado 3 ocorrer em $t \leq T$,

$f_{32}(t) \rightarrow$ é a função densidade da v.a. X_{32} com distribuição constante e período T,

$\delta(t - T) \rightarrow$ é a função impulso no ponto t-T, e

$f_1(t) * f_{32}(t) \rightarrow$ é a convolução de f_1 e f_{32}, a qual é expressa por $\int_0^t f_1(x) f_{32}(t-x)dx$,

é possível determinar $f_{32}(t)$ e, por extensão, P_2. Usando a propriedade da convolução, torna-se mais conveniente transpor a equação (A.4) para o domínio de Laplace, ou seja,

$$f_1^*(s).f_{32}^*(s) = e^{-sT} \quad , \tag{A.5}$$

de forma a obter a expressão de $f_{32}^*(s)$. Para tanto, determinemos primeiro $f_1^*(s)$:

$$f_1^*(s) = \int_0^\infty e^{-st} f_1(t)dt = \int_0^T e^{-st} \frac{P[t < X \leq t + \Delta t]}{P[X \leq T]} dt =$$

$$f_1^*(s) = \int_0^T e^{-st} \frac{\lambda_{13} e^{-\lambda_{13} t}}{1 - e^{-\lambda_{13} T}} dt = \frac{\lambda_{13}}{1 - e^{-\lambda_{13} T}} \int_0^T e^{-(s+\lambda_{13})t} dt$$

$$f_1^*(s) = \frac{\lambda_{13}}{1 - e^{-\lambda_{13} T}} \frac{1 - e^{-(s+\lambda_{13})T}}{s + \lambda_{13}} \quad . \tag{A.6}$$

Substituindo (A.6) em (A.5), tem-se:

$$\frac{\lambda_{13}}{1 - e^{-\lambda_{13} T}} \frac{1 - e^{-(s+\lambda_{13})T}}{s + \lambda_{13}} f_{32}^*(s) = e^{-sT} \quad ,$$

logo,

$$f_{32}^*(s) = \frac{e^{-sT}(s + \lambda_{13})(1 - e^{-\lambda_{13} T})}{\lambda_{13}(1 - e^{-(s+\lambda_{13})T})} \quad . \tag{A.7}$$

Substituindo (A.7) em (A.3), obtem-se:

$$P_2 e^{-sT} = f_{32}^*(s)|_{s=s+\lambda} = \frac{e^{-T(s+\lambda)}(s + \lambda + \lambda_{13})(1 - e^{-\lambda_{13} T})}{\lambda_{13}(1 - e^{-(s+\lambda+\lambda_{13})T})} \quad .$$

Fazendo s=0 e utilizando o valor de λ_{13} apresentado na figura A.1, ou seja, $\lambda_{13} = \lambda(1 - c_R)$, chega-se à expressão (A.8) de P_2.

$$P_2 = \frac{e^{-T\lambda}[\lambda(2 - c_R)][1 - e^{-\lambda(1-c_R)T}]}{\lambda(1 - c_R)[1 - e^{-\lambda(2-c_R)T}]} \quad . \tag{A.8}$$

Com o propósito de facilitar o processo de obtenção das probabilidades de ocupação dos estados, é possível suprimir o estado 3 do modelo embutido, resultando no modelo reduzido da figura A.2. Neste modelo, os estados 1, 2, 3 e 4 correspondem, respectivamente, aos estados 1, 2, 4 e 5 do modelo da figura A.1, e as probabilidades de transição entre os estados, a_{ij}, são dadas por:

$$a_{12} = \frac{c_A + c_R}{2} + \frac{1 - c_R}{2} P_2 \quad , \quad a_{13} = \frac{1 - c_R}{2}(1 - P_2) \quad , \quad a_{14} = \frac{1 - c_A}{2} \quad , \quad a_{21} = \frac{\mu}{\lambda + \mu}$$

$$a_{23} = \frac{\lambda}{\lambda + \mu} \quad , \quad a_{32} = 1 \quad , \quad a_{42} = 1$$

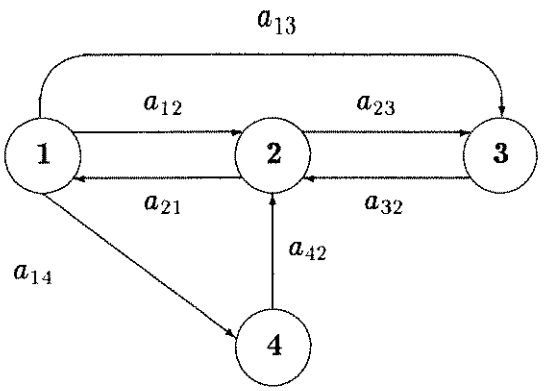


Figura A.2 - Modelo de Markov embutido,
referente ao modelo da figura A.1.

Cálculo da Indisponibilidade (I)

Para a obtenção das probabilidades estacionárias de ocupação dos estados 4 e 5 do modelo semi-markoviano da figura A.1, cuja soma exprime a indisponibilidade (I) do sistema, adotemos o procedimento apresentado em [HOW 71].

Seja ϕ_j a probabilidade estacionária de ocupação do estado j do modelo semi-markoviano, doravante chamada de probabilidade de estado, que é dada pela relação

$$\phi_j = \frac{\Pi_j \bar{\tau}_j}{\sum_{j=1}^{N} \Pi_j \bar{\tau}_j} = \frac{\Pi_j \bar{\tau}_j}{\bar{\tau}} \quad , \tag{A.9}$$

onde Π_j é a probabilidade do estado j no processo de Markov embutido, $\bar{\tau}_j$ é o tempo médio de espera no estado j, e N é o número de estados do modelo.

As probabilidades de estado do processo de markov embutido, por sua vez, são obtidas a partir da equação (A.10)

$$\vec{\Pi} = \vec{\Pi} P \quad , \tag{A.10}$$

onde $\vec{\Pi}$ é o vetor das probabilidades de estado e P é a matriz das probabilidades de transição do processo. A partir de (A.10), tem-se:

$$\vec{\Pi}[P - I] = 0 \quad .$$

Utilizando a condição $\sum_{j=1}^{N} \Pi_j = 1$ e substituindo a primeira coluna de $[P - I]$ por 1, resultando na matriz $[P - I]_m$, é possível escrever:

$$\vec{\Pi}[P - I]_m = [1\ 0 \cdots 0] \quad ,$$

logo,

$$\vec{\Pi} = [1\,0\cdots 0][P-I]_m^{-1} \quad . \tag{A.11}$$

Para o processo markoviano descrito pelo modelo da figura A.2, a matriz $[P-I]_m$ é dada por:

$$[P-I]_m = \begin{bmatrix} 1 & \frac{c_A+c_R}{2}+\frac{1-c_R}{2}P_2 & \frac{1-c_R}{2}(1-P_2) & \frac{1-c_A}{2} \\ 1 & -1 & \frac{\lambda}{\lambda+\mu} & 0 \\ 1 & 1 & -1 & 0 \\ 1 & 1 & 0 & -1 \end{bmatrix}$$

A solução de (A.11) leva às probabilidades Π_j , $j = 1,...,4$, as quais são apresentadas a seguir:

$\Pi_1 = (1-\frac{\lambda}{\lambda+\mu})D$,

$\Pi_2 = D$,

$\Pi_3 = \{\frac{\lambda}{\lambda+\mu}[\frac{1+c_R}{2}+\frac{1-c_R}{2}P_2+\frac{1-c_R}{2}(1-P_2)\}D$,

$\Pi_4 = [\frac{1-c_A}{2}-\frac{\lambda}{\lambda+\mu}\frac{(1-c_A)}{2}]D$,

onde

$$D = [3-\frac{c_A+c_R}{2}-\frac{1-c_R}{2}P_2-\frac{\lambda}{\lambda+\mu}[1-\frac{c_A+c_R}{2}-\frac{1-c_R}{2}P_2]^{-1} \quad .$$

A aplicação destas probabilidades, juntamente com os tempos médios de espera $\bar{\tau}_j$, $j = 1,...4$, ou seja,

$$\bar{\tau}_1 = \frac{1}{2\lambda} \quad , \quad \bar{\tau}_2 = \frac{1}{\lambda+\mu} \quad , \quad \bar{\tau}_3 = \frac{1}{\mu} \quad , \quad \bar{\tau}_4 = \frac{1}{\mu'}$$

à equação (A.12), torna possível a obtenção das probabilidades ϕ_j, em particular ϕ_3 e ϕ_4, uma vez que a indisponibilidade I é dada por:

$$\mathrm{I} = \phi_3 + \phi_4 \quad . \tag{A.12}$$

A tabela A.1 apresenta alguns valores de I em função de c_A, c_R e do período de supervisão em horas, ou seja, $T(h)$. Esta tabela permite a confrontação entre os valores de indisponibilidade obtidos através do modelo semi-markoviano da figura A.1, onde é adotada a hipótese de distribuição constante para o tempo de supervisão, e os valores obtidos através do modelo markoviano (modelo da figura 2.5), o qual considera uma distribuição exponencial para o tempo de supervisão. Estes valores são relacionados nas colunas I_{semi} e I_{markov}, respectivamente.

c_A	c_R	$T(h)$	I_{semi}	I_{markov}
0.9	0.9	24	2.53×10^{-6}	2.54×10^{-6}
0.9	0.9	1440	3.22×10^{-6}	3.73×10^{-6}
0.96	0.96	1440	1.3×10^{-6}	1.5×10^{-6}
0.9	0.9	10^4	6.75×10^{-6}	7.16×10^{-6}
0.99	0.99	10^4	6.9×10^{-7}	7.66×10^{-7}
0.999	0.999	10^4	8.7×10^{-8}	9.5×10^{-8}
0.99	0.9	10^4	4.5×10^{-6}	5.02×10^{-6}
0.9	0.9	10^5	1.25×10^{-5}	1.06×10^{-5}
0.9	0.9	10^7	1.25×10^{-5}	1.14×10^{-5}
0.99	0.99	10^5	1.27×10^{-6}	1.17×10^{-6}

Tabela A.1 - Indisponibilidade de um sistema *hot standby* 1:1 : valores obtidos para $\mu = 1$, $\mu' = 4$ e $\lambda = 10^{-4}$ $falhas/h$.

Através dos resultados apresentados na tabela A.1, é possível notar que:

- Os valores de indisponibilidade obtidos com os modelos markoviano e semi-markoviano praticamente não apresentam diferenças, principalmente quando $T(h) << \lambda^{-1}$, o que era esperado. Quando $T(h) \rightarrow \lambda^{-1}$ e $T(h) > \lambda^{-1}$, os resultados obtidos não apresentam diferenças significativas.

- Para $T(h) \leq \lambda^{-1}$, os valores de indisponibilidade obtidos com o modelo markoviano apresentam-se mais conservativos do que os obtidos através do modelo semi-markoviano.

Com isto é possível concluir que a hipótese de distribuição exponencial para o tempo de supervisão, e a consequente utilização de um modelo markoviano, é adequada a tal tipo de avaliação, mesmo em se tratando de períodos de supervisão da ordem de λ^{-1}. Além disso, a adoção desta hipótese é reforçada pela maior simplicidade apresentada pelos modelos de Markov, no que diz respeito ao processo de obtenção das probabilidades de ocupação dos estados.