

UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS E TECNOLOGIA
COORDENAÇÃO DE PÓS-GRADUAÇÃO EM INFORMÁTICA

AVALIAÇÃO DE DISPONIBILIDADE DE REDES DE COMPUTADORES
BASEADAS NA ARQUITETURA CLIENTE/SERVIDOR EM N-CAMADAS

FLÁVIA ESTÉLIA SILVA COELHO

CAMPINA GRANDE, SETEMBRO DE 2000.

UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS E TECNOLOGIA
COORDENAÇÃO DE PÓS-GRADUAÇÃO EM INFORMÁTICA

FLÁVIA ESTÉLIA SILVA COELHO

AVALIAÇÃO DE DISPONIBILIDADE DE REDES DE COMPUTADORES
BASEADAS NA ARQUITETURA CLIENTE/SERVIDOR EM N-CAMADAS

Dissertação apresentada ao curso de Mestrado
em Informática da Universidade Federal da
Paraíba, em cumprimento às exigências
parciais para a obtenção do grau de Mestre.

ORIENTADOR: JACQUES PHILIPPE SAUVÉ
ÁREA DE CONCENTRAÇÃO: CIÊNCIA DA COMPUTAÇÃO
LINHA DE PESQUISA: REDES DE COMPUTADORES

C672a Coelho, Flavia Estelia Silva
Avaliacao de disponibilidade de redes de computadores baseadas na arquitetura cliente/servidor em n-camadas / Flavia Estelia Silva Coelho. - Campina Grande, 2000. 127 f.

Dissertacao (Mestrado em Informatica) - Universidade Federal da Paraiba, Centro de Ciencias e Tecnologia.

1. Redes de Computadores 2. Disponibilidade 3. Avaliacao de Disponibilidade 4. Dissertacao - Informatica I. Sauve, Jacques Philippe II. Universidade Federal da Paraiba - Campina Grande (PB)

CDU 004.7(043)

**AVALIAÇÃO DE DISPONIBILIDADE DE REDES DE
COMPUTADORES: UM ESTUDO DE CASO PARA A
ARQUITETURA CLIENTE/SERVIDOR EM N-CAMADAS**

FLÁVIA ESTÉLIA SILVA COELHO

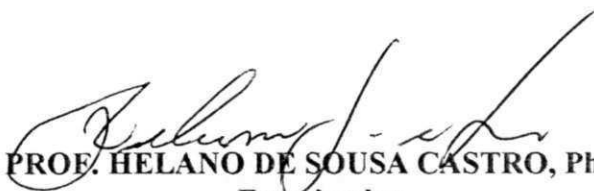
DISSERTAÇÃO APROVADA EM 29.08.2000



PROF. JACQUES PHILIPPE SAUVÉ, Ph.D
Orientador



PROF. JOSÉ ANTÃO BELTRÃO MOURA, Ph.D
Examinador



PROF. HELANO DE SOUSA CASTRO, Ph.D
Examinador

CAMPINA GRANDE – PB.

AGRADECIMENTOS

É muito gratificante terminar mais uma etapa em nossa vida e poder reconhecer a ajuda, a compreensão e o carinho de tantas pessoas. Melhor ainda, é poder dar-lhes sinceros agradecimentos.

Agradecimentos mais que especiais aos meus pais e melhores amigos - Fátima e Espedito. Obrigada pelo amor, amizade e respeito sinceros; por toda a dedicação, compromisso e carinho com que me educaram. Tenho em vocês um *exemplo maior* de dignidade e luta.

Sinceros agradecimentos ao meu orientador Jacques Sauvé, pelo empenho, compreensão e seriedade no decorrer de todo este trabalho. Com sua força e amor ao trabalho, aprendi que o trabalho nos dignifica e que vale a pena nos dedicar ao que estamos fazendo, pois será sempre válido, apesar de todas as dificuldades.

Agradecimentos aos meus manos Fábio (agradecimento especial a ti, incentivador e amigo leal, em todos os momentos. Sei que sempre poderei contar contigo) e Júnior, a minha sobrinha Geovanna e ao meu tio Epifânio, pelos incentivos e carinho.

Agradeço também a todos os meus *dream* amigos (em especial a Livia, Ana Karla, Natasha, Raissa e Guga); aos amigos do Laboratório de Redes de Computadores (especialmente a Peter, Karina, Giovanni, Juliana e Marcus Vinicius) e a meus demais amigos, por todo o carinho e amizade.

Agradecimentos especiais a todos os professores e funcionários do Departamento de Sistemas e Computação, em especial a Lilian, Aninha e Manuela.

RESUMO

Atualmente, redes de computadores corporativas são baseadas na arquitetura cliente/servidor em 3-camadas (ou n -camadas). Para que clientes possam receber serviços completos de tais redes, algumas restrições quanto à conectividade entre os clientes e servidores (e entre servidores) devem ser atendidas. No entanto, tais conectividades podem ser afetadas por falhas em enlaces, equipamentos de interconexão e nos próprios servidores da rede. Sendo assim, falhas podem afetar a disponibilidade serviços para determinados clientes, sem afetar os demais clientes. Considerando esse contexto, este trabalho propõe uma nova medida de disponibilidade que expressa a fração de clientes que estão recebendo serviços no contexto de redes de computadores baseadas na arquitetura cliente/servidor em n -camadas, considerando possíveis falhas em seus componentes. Adicionalmente, propõe-se um método de avaliação de disponibilidade que utiliza a nova medida proposta, ao avaliar a disponibilidade de redes de computadores baseadas na arquitetura cliente/servidor em ou n -camadas.

ABSTRACT

Nowadays, corporate computer networks are based on multi-tier client/server architectures. Some constraints on the connectivity among clients and servers (and among servers) must be satisfied for the clients to receive complete services from the network. However, failures in the links, interconnection equipments and servers can affect this connectivity. Therefore, failures can affect service availability for some clients, but not all. This work proposes a new network availability measure: the fraction of clients that are receiving complete services in computer networks based on multi-tier client/server architectures, considering its component failures. In addition, a method for the evaluation of network availability using the new measure is proposed.

SUMÁRIO

1. INTRODUÇÃO	1
1.1. MOTIVAÇÃO E RELEVÂNCIA	1
1.2. OBJETIVOS	4
1.2.1. Objetivos Gerais	4
1.2.2. Objetivos Específicos	4
1.3. ORGANIZAÇÃO	5
2. DISPONIBILIDADE DE REDES DE COMPUTADORES	6
2.1. UMA VISÃO GERAL SOBRE A DISPONIBILIDADE DE REDES DE COMPUTADORES	6
2.2. MÉTODOS DE AVALIAÇÃO DE DISPONIBILIDADE DE REDES DE COMPUTADORES	12
3. UMA NOVA MEDIDA DE DISPONIBILIDADE DE REDES DE COMPUTADORES	17
3.1. AVALIAÇÃO DE DISPONIBILIDADE DE REDES DE COMPUTADORES BASEADAS NA ARQUITETURA CLIENTE/SERVIDOR EM N-CAMADAS	18
3.2. ESCOPO CONSIDERADO	21
3.3. DEFINIÇÃO FORMAL DA NOVA MEDIDA DE DISPONIBILIDADE	24
4. UM MÉTODO DE AVALIAÇÃO DE DISPONIBILIDADE DE REDES DE COMPUTADORES BASEADAS NA ARQUITETURA CLIENTE/SERVIDOR EM N-CAMADAS	38
4.1. MODELO DE DISPONIBILIDADE	38
4.2. CONSIDERAÇÕES PRELIMINARES PARA A DEFINIÇÃO FORMAL	39
4.3. DEFINIÇÃO FORMAL DO MÉTODO DE AVALIAÇÃO DE DISPONIBILIDADE DE REDES DE COMPUTADORES BASEADAS NA ARQUITETURA CLIENTE/SERVIDOR EM N-CAMADAS	44
4.4. COMPLEXIDADE COMPUTACIONAL	55
5. ESTUDOS DE CASO	58
5.1. CONSIDERAÇÕES PARA OS ESTUDOS DE CASO	59
5.2. ESTUDO DO COMPORTAMENTO DOS LIMITES SUPERIOR E INFERIOR DE DISPONIBILIDADE EM UMA REDE DE COMPUTADORES	60
5.3. ESTUDO DO IMPACTO DA ADIÇÃO DE COMPONENTES EM UMA REDE DE COMPUTADORES	63
5.3.1. Adição de Clientes	63
5.3.2. Adição de Servidores Departamentais e Servidores Corporativos	65
5.4. ESTUDO DO IMPACTO DA VARIAÇÃO DA DISPONIBILIDADE DOS COMPONENTES DE UMA REDE DE COMPUTADORES	66
5.4.1. Variando a Disponibilidade dos Componentes	66

5.4.2. <i>Variando a Disponibilidade dos Enlaces Individuais</i>	68
5.5. ESTUDO DO IMPACTO DA INSERÇÃO DE REDUNDÂNCIA DE RECURSOS FÍSICOS EM UMA REDE DE COMPUTADORES	70
5.6. ESTRUTURA <i>COLLAPSED BACKBONE VERSUS</i> ESTRUTURA HIERÁRQUICA	73
5.7. ESTUDO DO COMPORTAMENTO DO TEMPO DE PROCESSAMENTO DO MÉTODO DE AVALIAÇÃO DE DISPONIBILIDADE	74
6. CONCLUSÃO	79
6.1. CONCLUSÕES	79
6.2. CONTRIBUIÇÕES	81
6.3. TRABALHOS FUTUROS	83
REFERÊNCIAS BIBLIOGRÁFICAS	85
APÊNDICE A - TEORIA DA CONFIABILIDADE	94
A.1. CONCEITOS BÁSICOS	94
A.2. CONFIABILIDADE	95
A.2.1. <i>Sistemas Computacionais Organizados em Série</i>	97
A.2.2. <i>Sistemas Computacionais Organizados em Paralelo</i>	97
A.2.3. <i>Sistemas Computacionais Organizados em Série e em Paralelo</i>	98
A.3. TEMPO MÉDIO PARA FALHAR - MTTR E TEMPO MÉDIO ENTRE FALHAS - MTBF	99
A.4. TEMPO MÉDIO DE REPARO - MTTR	100
A.5. DISPONIBILIDADE	100
APÊNDICE B - ARQUITETURA CLIENTE/SERVIDOR EM <i>N</i> -CAMADAS	102
B.1. ARQUITETURA CLIENTE/SERVIDOR EM 2-CAMADAS	102
B.2. ARQUITETURA CLIENTE/SERVIDOR EM <i>N</i> -CAMADAS	103
APÊNDICE C - CONFIGURAÇÕES TOPOLÓGICAS UTILIZADAS NOS ESTUDOS DE CASO	105

LISTA DE FIGURAS

FIGURA 3.1 - ARQUITETURA CLIENTE/SERVIDOR EM n -CAMADAS	19
FIGURA 3.2 - REDE DE COMPUTADORES EXEMPLO	23
FIGURA 3.3 - ESTRUTURA DA MATRIZ DE ADJACÊNCIAS ORIGINAIS	25
FIGURA 3.4 - ESTRUTURA DA MATRIZ DE CONECTIVIDADES DESEJADAS	28
FIGURA 5.1 - LIMITES SUPERIOR E INFERIOR DA NOVA MEDIDA DE DISPONIBILIDADE X NÚMERO DE ESTADOS DE FALHA MAIS PROVÁVEIS DA REDE	62
FIGURA 5.2 - LIMITES SUPERIOR E INFERIOR DA NOVA MEDIDA DE DISPONIBILIDADE X NÚMERO DE CLIENTES	64
FIGURA 5.3 - LIMITES SUPERIOR E INFERIOR DA NOVA MEDIDA DE DISPONIBILIDADE X NÚMERO DE SERVIDORES DEPARTAMENTAIS E CORPORATIVOS	65
FIGURA 5.4 - LIMITES SUPERIOR E INFERIOR DA NOVA MEDIDA DE DISPONIBILIDADE X DISPONIBILIDADES DOS COMPONENTES	67
FIGURA 5.5 - LIMITES SUPERIOR E INFERIOR DA NOVA MEDIDA DE DISPONIBILIDADE X DISPONIBILIDADES DOS ENLACES	70
FIGURA 5.6 - TEMPO DE PROCESSAMENTO X NÚMERO TOTAL DE COMPONENTES DA REDE	75
FIGURA 5.7 - TEMPO DE PROCESSAMENTO X NÚMERO DE ESTADOS DE FALHA MAIS PROVÁVEIS DA REDE	77
FIGURA A.1 - SISTEMA COMPUTACIONAL ORGANIZADO EM SÉRIE	97
FIGURA A.2 - SISTEMA COMPUTACIONAL ORGANIZADO EM PARALELO	98
FIGURA A.3 - SISTEMA COMPUTACIONAL ORGANIZADO EM SÉRIE E EM PARALELO	98
FIGURA A.4 - CICLO DE VIDA DE UM SISTEMA COMPUTACIONAL	99
FIGURA A.5 - ESTADOS DE UM SISTEMA COMPUTACIONAL	100
FIGURA B.1 - ARQUITETURA CLIENTE/SERVIDOR EM 2-CAMADAS	102
FIGURA B.2 - ARQUITETURA CLIENTE/SERVIDOR EM n -CAMADAS	103
FIGURA C.1 - CONFIGURAÇÃO TOPOLÓGICA 1	105
FIGURA C.2 - CONFIGURAÇÃO TOPOLÓGICA 2	106

FIGURA C.3 - CONFIGURAÇÃO TOPOLÓGICA 3	107
FIGURA C.4 - CONFIGURAÇÃO TOPOLÓGICA 4	108
FIGURA C.5 - CONFIGURAÇÃO TOPOLÓGICA 5	109
FIGURA C.6 - CONFIGURAÇÃO TOPOLÓGICA 6	110
FIGURA C.7 - CONFIGURAÇÃO TOPOLÓGICA 7	111
FIGURA C.8 - CONFIGURAÇÃO TOPOLÓGICA 8	112
FIGURA C.9 - CONFIGURAÇÃO TOPOLÓGICA 9	113
FIGURA C.10 - CONFIGURAÇÃO TOPOLÓGICA 10	114
FIGURA C.11 - CONFIGURAÇÃO TOPOLÓGICA 11	115
FIGURA C.12 - CONFIGURAÇÃO TOPOLÓGICA 12	116
FIGURA C.13 - CONFIGURAÇÃO TOPOLÓGICA 13	117
FIGURA C.14 - CONFIGURAÇÃO TOPOLÓGICA 14	118
FIGURA C.15 - CONFIGURAÇÃO TOPOLÓGICA 15	119
FIGURA C.16 - CONFIGURAÇÃO TOPOLÓGICA 16	120
FIGURA C.17 - CONFIGURAÇÃO TOPOLÓGICA 17	121
FIGURA C.18 - CONFIGURAÇÃO TOPOLÓGICA 18	122
FIGURA C.19 - CONFIGURAÇÃO TOPOLÓGICA 19	123
FIGURA C.20 - CONFIGURAÇÃO TOPOLÓGICA 20	124
FIGURA C.21 - CONFIGURAÇÃO TOPOLÓGICA 21	125
FIGURA C.22 - CONFIGURAÇÃO TOPOLÓGICA 22	126
FIGURA C.23 - CONFIGURAÇÃO TOPOLÓGICA 23	127

LISTA DE TABELAS

TABELA 2.1 - DISPONIBILIDADE EM RELAÇÃO AOS TEMPOS DE QUEDA POR PERÍODOS ESPECÍFICOS	8
TABELA 3.1 - MATRIZ DE ADJACÊNCIAS ORIGINAIS DA REDE-EXEMPLO 3.2	27
TABELA 3.2 - DEFINIÇÃO DO VALOR DE MCD'_{JK}	29
TABELA 3.3 - MATRIZ DE CONECTIVIDADES DESEJADAS PARA O CLIENTE 25 DA REDE-EXEMPLO 3.2	30
TABELA 3.4 - MATRIZ DE ADJACÊNCIAS PARA O ESTADO DE FALHA E DA REDE-EXEMPLO 3.2	32
TABELA 3.5 - MATRIZ DE CONECTIVIDADES REAIS PARA O ESTADO DE FALHA E DA REDE 3.2	34
TABELA 4.1 - COMPLEXIDADE COMPUTACIONAL DAS VÁRIAS ETAPAS DO MÉTODO DE AVALIAÇÃO DE DISPONIBILIDADE DE REDES DE COMPUTADORES BASEADAS NA ARQUITETURA CLIENTE/SERVIDOR EM N-CAMADAS	56
TABELA 5.1 - DISPONIBILIDADES DOS COMPONENTES DAS CONFIGURAÇÕES TOPOLÓGICAS UTILIZADAS NOS ESTUDOS DE CASO	59
TABELA 5.2 - NÚMERO DE ESTADOS DE FALHA MAIS PROVÁVEIS DA REDE EM RELAÇÃO AOS LIMITES SUPERIOR E INFERIOR DA NOVA MEDIDA DE DISPONIBILIDADE	61
TABELA 5.3 - NÚMERO DE CLIENTES EM RELAÇÃO AOS LIMITES SUPERIOR E INFERIOR DA NOVA MEDIDA DE DISPONIBILIDADE	63
TABELA 5.4 - NÚMERO DE SERVIDORES DEPARTAMENTAIS E CORPORATIVOS EM RELAÇÃO AOS LIMITES SUPERIOR E INFERIOR DA NOVA MEDIDA DE DISPONIBILIDADE ..	65
TABELA 5.5 - LIMITES SUPERIOR E INFERIOR DA NOVA MEDIDA DE DISPONIBILIDADE EM RELAÇÃO À VARIAÇÃO DAS DISPONIBILIDADES DOS COMPONENTES DA REDE	67
TABELA 5.6 - DISPONIBILIDADE EM RELAÇÃO AOS TEMPOS DE QUEDA POR PERÍODOS ESPECÍFICOS	68

TABELA 5.7 - DISPONIBILIDADES CONSIDERADAS PARA ESTUDO DE CASO ESPECÍFICO	69
TABELA 5.8 - LIMITES SUPERIOR E INFERIOR DA NOVA MEDIDA DE DISPONIBILIDADE EM RELAÇÃO À VARIAÇÃO DAS DISPONIBILIDADES DOS ENLACES INDIVIDUAIS DA REDE	69
TABELA 5.9 - LIMITES SUPERIOR E INFERIOR DA NOVA MEDIDA DE DISPONIBILIDADE EM RELAÇÃO AO NÚMERO DE ENLACES INDIVIDUAIS DA REDE	71
TABELA 5.10 - DISPONIBILIDADES EM RELAÇÃO AOS TEMPOS DE QUEDA POR PERÍODOS ESPECÍFICOS	71
TABELA 5.11 - LIMITES SUPERIOR E INFERIOR DA NOVA MEDIDA DE DISPONIBILIDADE EM RELAÇÃO A MUDANÇAS NA ESTRUTURA TOPOLÓGICA DA REDE	73
TABELA 5.12 - TEMPO DE PROCESSAMENTO EM RELAÇÃO AO TAMANHO DA REDE	75
TABELA 5.13 - TEMPO DE PROCESSAMENTO EM RELAÇÃO AO NÚMERO DE ESTADOS DE FALHA MAIS PROVÁVEIS DA REDE	76
TABELA 6.1 - COMPARAÇÃO ENTRE MÉTODOS DE AVALIAÇÃO DE DISPONIBILIDADE DE REDES DE COMPUTADORES	81

CAPÍTULO 1

INTRODUÇÃO

Neste capítulo, apresentam-se a motivação, a relevância, os objetivos gerais e específicos deste trabalho de dissertação e, por fim, apresenta-se a estrutura organizacional do restante dos capítulos que compreendem este trabalho.

1.1. Motivação e Relevância

Atualmente, redes de computadores corporativas são baseadas na arquitetura cliente/servidor em 3-camadas (ou n -camadas), ou seja, são logicamente divididas em três módulos: camada de serviços dos usuários (clientes), camada intermediária (servidores corporativos de aplicação, de correio eletrônico, etc.) e camada de dados (servidores corporativos de bancos de dados).

Redes de computadores corporativas têm muitos clientes e servidores. Os servidores provêm serviços (serviços tais como acesso a Web e armazenamento de arquivos) necessários às aplicações típicas da rede para uma parte (no caso de servidores departamentais) ou para todos os usuários da rede (no caso de servidores corporativos). Exemplos de aplicações típicas de redes corporativas são aplicações baseadas em comércio eletrônico e sistemas de colaboração entre grupos de trabalho (conjuntos de clientes específicos).

Redes de computadores devem atender aos requisitos das aplicações que suportam, ou seja, devem prover uma infra-estrutura que atenda aos requisitos específicos impostos por tais aplicações.

Este trabalho considera as aplicações que requerem alta disponibilidade para executarem suas funcionalidades desejadas, a fim de satisfazer os clientes e as próprias empresas, como um todo. Exemplos de tais aplicações são sistemas baseados em comércio eletrônico, sistemas de transações bancárias, sistemas de acompanhamento de linha de produção industrial, entre outras aplicações [3COM98, BUCHANAN-WEAL99].

Em redes corporativas baseadas na arquitetura cliente/servidor em n -camadas, para que os clientes recebam os serviços disponibilizados pelos servidores, as conectividades entre os clientes e servidores (e entre servidores) são consideradas críticas e devem satisfazer algumas restrições.

Para um cliente particular receber serviços completos da rede, ele deve acessar os seus respectivos servidores departamentais e os servidores corporativos da camada intermediária e estes, por sua vez, devem acessar uns aos outros e os servidores corporativos da camada de dados. Estes últimos não precisam acessar uns aos outros, precisam apenas acessar os servidores corporativos da camada intermediária.

Entretanto, redes de computadores são constituídas por componentes e estes, por sua vez, são suscetíveis a falhas. Sendo assim, falhas em enlaces que conectam clientes e servidores (e servidores entre si), nos próprios servidores e em equipamentos de interconexão podem afetar, de forma crítica, a disponibilidade de serviços para os clientes da rede. Estabelecer e manter uma determinada rede operacional e prover, satisfatoriamente, serviços aos clientes têm implicações críticas para as empresas, uma vez que falhas podem afetar seus lucros, custos e produtividade, além de causar insatisfação aos clientes da empresa.

No entanto, falhas em determinados componentes da rede podem afetá-la apenas parcialmente, isto é, a rede pode estar parcialmente operacional e os clientes são afetados apenas parcialmente, ou seja, uma parte dos clientes continuará a receber os serviços da rede, apesar das falhas.

Considerando esse contexto, faz-se necessário medir a quantidade de clientes de redes de computadores baseadas na arquitetura cliente/servidor em n -camadas que estão recebendo serviços completos da rede considerando, para tanto, as possíveis ocorrências de falhas nos componentes da rede.

Medidas de disponibilidade existentes [AGGARWAL ET AL.75, BALL79, BALL-VAN SLYKE77, BAUER ET AL.87, HÄNSLER ET AL.74, KE-WANG97, KUBAT89, MYRVOLD96, VAN SLYKE-FRANK72, VAN SLYKE ET AL.75, TORRIERI94], embora sejam baseadas na conectividade parcial da rede, não expressam as reais necessidades quanto às conectividades parciais entre os clientes e servidores (e entre servidores) para que os clientes recebam serviços no contexto de redes de computadores baseadas na arquitetura cliente/servidor em n -camadas.

Sendo assim, este trabalho tem por objetivos: 1) propor formalmente uma nova medida de disponibilidade de redes de computadores baseadas na arquitetura cliente/servidor em n -camadas, isto é, a fração de clientes que estão recebendo serviços completos da rede, considerando possíveis falhas em seus componentes e 2) propor um método de avaliação de disponibilidade para o cálculo de tal medida.

Métodos de avaliação de disponibilidade auxiliam os projetistas nos processos de projetos de redes de computadores com requisitos específicos de disponibilidade, uma vez que a disponibilidade total da rede pode ser avaliada utilizando uma medida específica de disponibilidade e considerando os efeitos das possíveis falhas nos componentes da rede.

Tais métodos de avaliação de disponibilidade permitem que topologias, tecnologias de redes, estratégias de interconexão e mecanismos que aumentam a disponibilidade de redes de computadores sejam analisados, a fim de encontrar uma configuração que garanta o valor de disponibilidade requisitado, levando em consideração restrições para o projeto em questão (minimizar custos ou maximizar desempenho, por exemplo).

Por exemplo, muitas empresas preferem empregar caminhos redundantes (caminhos alternativos para roteamento) apenas nas porções da rede que carregam informação crítica (em termos de disponibilidade), uma vez que uma topologia totalmente redundante, na maioria dos casos, tem custo proibitivo.

Vale ressaltar que o método de avaliação de disponibilidade proposto neste trabalho, realiza a avaliação analisando o efeito de possíveis falhas no contexto topológico da rede, isto é, não são levadas em consideração as influências que estratégias ou mecanismos que aumentam a disponibilidade de redes de computadores podem causar à disponibilidade total das redes analisadas.

1.2. Objetivos

1.2.1. Objetivos Gerais

- 1) Propor uma nova medida de disponibilidade de redes de computadores baseadas na arquitetura cliente/servidor em 3-camadas (ou n -camadas);
- 2) Propor um método de avaliação de disponibilidade que utilize a nova medida de disponibilidade proposta.

1.2.2. Objetivos Específicos

- 1) Fazer um levantamento das medidas de disponibilidade de redes de computadores existentes;
- 2) Verificar se algumas das medidas de disponibilidade existentes atendem às restrições de conectividade entre clientes e servidores (e entre servidores), necessárias para que clientes recebam serviços completos de redes de computadores baseadas na arquitetura cliente/servidor em 3-camadas (ou n -camadas);
- 3) Definir formalmente uma nova medida de disponibilidade de redes de computadores baseadas na arquitetura cliente/servidor em 3-camadas (ou n -camadas);
- 4) Definir formalmente o método de avaliação de disponibilidade de redes de computadores baseadas na arquitetura cliente/servidor em 3-camadas (n -camadas), que utilize a nova medida de disponibilidade proposta;
- 5) Realizar estudos de caso, a fim de analisar os resultados experimentais obtidos, ao empregar o método de avaliação de disponibilidade de redes proposto para avaliar a disponibilidade de configurações topológicas particulares, considerando a nova medida de disponibilidade de redes de computadores baseadas na arquitetura cliente/servidor em 3-camadas (ou n -camadas).

1.3. Organização

Os capítulos seguintes deste trabalho estão estruturados segundo a ordem descrita a seguir.

No capítulo 2, apresenta-se uma visão geral a respeito de disponibilidade de redes de computadores. Para tanto, são apresentadas noções básicas a respeito de confiabilidade, disponibilidade, tempo médio para falhar, tempo médio entre falhas e tempo médio para reparo e, por fim, são descritos os efeitos de falhas em componentes para a disponibilidade total da rede. Em seguida, descrevem-se os principais métodos existentes de avaliação de disponibilidade de redes de computadores, descrevendo assim, os modelos, medidas de disponibilidade e técnicas de análise tipicamente utilizados por tais métodos de avaliação.

Uma nova medida de disponibilidade de redes de computadores baseadas na arquitetura cliente/servidor em 3-camadas (ou n -camadas) é definida formalmente, no capítulo 3. Para tanto, apresenta-se o contexto de redes baseadas na arquitetura cliente/servidor em 3-camadas (ou n -camadas) e o escopo considerado ao definir formalmente a nova medida.

O método de avaliação de disponibilidade de redes que utiliza a nova medida de disponibilidade proposta é definido no capítulo 4. Para tanto, apresentam-se o modelo considerado, a descrição formal e a complexidade computacional do método de avaliação de disponibilidade proposto.

No capítulo 5, são apresentados alguns estudos de caso com a finalidade de exemplificar o emprego do método de avaliação de disponibilidade de redes e caracterizar o comportamento da nova medida de disponibilidade proposta, no contexto de configurações topológicas particulares.

Informações adicionais a respeito da teoria da confiabilidade e da arquitetura cliente/servidor em 3-camadas (ou n -camadas) são disponibilizadas nos apêndices A e B, respectivamente. Finalmente, as configurações topológicas utilizadas ao realizar os estudos de caso, apresentados no capítulo 5 podem ser consultadas no apêndice C.

CAPÍTULO 2

DISPONIBILIDADE DE REDES DE COMPUTADORES

Atualmente, a disponibilidade é um dos principais requisitos impostos por aplicações às redes de computadores [3COM98, BUCHANAN-WEAL99, MCCABE98]. Sendo assim, projetistas devem considerar esta questão e propor meios de, utilizando as estratégias e mecanismos existentes para aumentar disponibilidade, garantir que uma determinada rede de computadores atinja um valor pré-especificado de disponibilidade.

Para auxiliar os projetistas de redes a executar tal tarefa, há métodos de avaliação de disponibilidade de redes que permitem avaliar a disponibilidade total obtida considerando, para tanto, as possíveis falhas em componentes de uma rede de computadores específica.

Neste capítulo, é dada uma visão geral a respeito da disponibilidade de redes de computadores. Em seguida, os modelos tradicionais, as medidas de disponibilidade utilizadas tipicamente e a complexidade computacional dos principais métodos de avaliação de disponibilidade de redes de computadores são discutidos.

Vale ressaltar, inicialmente, que os métodos de avaliação de disponibilidade existentes [AGGARWAL ET AL.75, AGGARWAL-RAI81, BALL79, BALL-VAN SLYKE77, HÄNSLER ET AL.74, KERSHENBAUM-VAN SLYKE73, KUBAT89] utilizam o termo confiabilidade para expressar disponibilidade de redes de computadores. No entanto, é válido salientar que as análises realizadas por tais métodos de avaliação consideram a disponibilidade dos componentes da rede, ao invés da confiabilidade de tais componente.

2.1. Uma Visão Geral sobre a Disponibilidade de Redes de Computadores

Redes de computadores são constituídas por estações de trabalho, servidores,

equipamentos de interconexão, entre outros dispositivos, e enlaces que os interconectam. Cada um desses componentes pode ser considerado um sistema constituído por elementos com características próprias de falha. A teoria da confiabilidade¹ compreende os conceitos relevantes ao estudo do comportamento de falhas em tais sistemas.

A confiabilidade de um sistema é definida como a sua probabilidade de estar operacional continuamente, durante um intervalo específico $[0, t]$, dado que ele estava operacional em $t = 0$ [BLAKE79].

Observe que a definição de confiabilidade de sistemas não considera os tempos necessários para executar reparos ou processos de recuperação, que são responsáveis por restaurar o estado operacional de sistemas no evento de falhas. Na realidade, apenas o tempo de operação contínua, sem interrupções, é considerado. Em outras palavras, confiabilidade está diretamente relacionada ao tempo médio para falhar (MTTF - *Mean Time To Failure*) [BLAKE79, SIEWIOREK-SWARZ92].

No entanto, quando sistemas ou seus elementos constituintes falham, normalmente, são reparados ou trocados por novos sistemas ou elementos, respectivamente. Pode-se, então, assumir que o tempo de vida de um sistema ou elemento particular compreende períodos alternados de operação e falha.

Considerando esse comportamento, pode-se medir os seguintes valores de recuperabilidade de sistemas: MTBF (*Mean Time Between Failures*) - tempo médio entre falhas e MTTR (*Mean Time To Repair*) - tempo médio para reparo [BLAKE79], onde $MTBF = MTTF + MTTR$ [SIEWIOREK-SWARZ92].

Em resumo, o tempo médio entre falhas refere-se ao tempo transcorrido entre uma falha e outra (incluindo o tempo necessário para reparo/restauração) enquanto o tempo médio para reparo engloba o intervalo que inicia no instante em que uma falha é detectada até o instante em que o sistema ou seus elementos constituintes sejam completamente reparados ou trocados.

A disponibilidade de um sistema é a sua probabilidade de estar operacional no instante de tempo t [BLAKE79].

Através do MTTF e MTTR, pode-se definir a disponibilidade de sistemas como uma função da frequência de ocorrência de falhas e do tempo necessário para seu reparo ou

¹ Para obter maiores informações a respeito da teoria da confiabilidade, vide apêndice A.

troca, isto é, a disponibilidade é uma função direta do MTTF e MTTR (especificamente, disponibilidade é $MTTF/(MTTF+MTTR)$) ou disponibilidade é tempo operando/(tempo operando + tempo parado) - vide apêndice A, para maiores detalhes. Por exemplo, um MTTF de 8000 horas e um MTTR de 4 horas expressam uma disponibilidade de 99.95%.

A tabela 2.1 apresenta alguns valores de disponibilidade, em termos dos tempos de queda por períodos de tempo específicos.

<i>Disponibilidade</i>	<i>Tempo de Queda</i>		
	<i>Mensal</i>	<i>Semanal</i>	<i>Diário</i>
95%	36.5 h	8.4 h	1.2 h
99%	7.3 h	1.68 h	14.4 min
99.5%	3.7 h	50.5 min	7.2 min
99.9%	43.8 min	10.08 min	1.44 min
99.95%	21.9 min	5.05 min	43.2 s
99.98%	8.75 min	2.0 min	17.3 s
99.99%	4.4 min	1.0 min	8.7 s
99.999%	26.28 s	6.048 s	0.864 s

Tabela 2.1 - Disponibilidade em relação aos Tempos de Queda por Períodos Específicos

Na prática, os dispositivos e enlaces que constituem uma rede de computadores têm disponibilidades que variam, geralmente, entre 99.9 e 99.999% [3COM97A, 3COM97B, 3COM00, CISCO99C, CISCO00, MAGALLANES98].

Segundo [MCCABE98], disponibilidades em torno de 95% são adequadas apenas para testes e protótipos, sendo assim são valores inaceitáveis para sistemas reais. A maioria dos sistemas, na prática, opera com disponibilidades acima de 99.9%. Disponibilidades a partir de 99.98% indicam alta disponibilidade e são recomendadas para sistemas computacionais cujas falhas são críticas para o estado operacional da rede.

É interessante ressaltar que confiabilidade e disponibilidade são medidas do tempo em que um sistema está operacional. Entretanto, a disponibilidade leva em consideração as interrupções necessárias para efetuar processos de reparo, troca ou recuperação enquanto

confiabilidade não leva em conta tais interrupções.

Considerando toda a discussão anterior, afirma-se que redes de computadores podem conter partes não operacionais devido a falhas em seus componentes² e estes, por sua vez, podem estar não operacionais devido a um dos tipos de queda: quedas planejadas ou quedas não planejadas.

Quedas planejadas são devidas a ações de manutenção, tais como atualizações em hardware e software, inicializações, entre outros processos. Tempos de quedas planejadas podem ser reduzidos empregando estratégias tais como: realizar manutenção enquanto o sistema está operacional e reduzir o tempo gasto para executar tarefas que podem ser efetuadas apenas quando o sistema não está operacional.

Quedas não planejadas ocorrem devido a falhas em hardware e software e a influências ambientais, tais como desastres naturais e erros humanos, entre outros fatores. As seguintes estratégias podem ser utilizadas para reduzir o tempo de quedas não planejadas: reduzir o tempo gasto em processos de recuperação e evitar pontos únicos de falha, isto é, utilizar sistemas e elementos constituintes redundantes. Redundância é um mecanismo utilizado para aumentar a confiabilidade de redes de computadores, sistemas computacionais, etc. e será discutida adiante.

O tempo de duração de quedas, sejam planejadas ou não planejadas, representa o tempo em que um sistema não está operacional, isto é, não está disponível para executar sua funcionalidade própria.

Falhas nos componentes de uma rede de computadores afetam a sua disponibilidade de três maneiras, basicamente:

- A falha de um dispositivo pode afetar a comunicação entre ele e o resto da rede. Por exemplo, caso um servidor de uma rede corporativa falhe, nenhum de seus clientes terá acesso a seus serviços;
- Falhas em dispositivos, que estão no caminho entre outros dispositivos, afetam a comunicação entre eles. Por exemplo, falhas em equipamentos de interconexão afetam a comunicação entre as sub-redes que eles interconectam;
- Falhas em enlaces também afetam a comunicação entre os dispositivos da rede.

Há aplicações que requerem redes de computadores provendo uma infra-estrutura de disponibilidade específica ou de alta disponibilidade (níveis acima de 99.98% de disponibilidade), que seja suficiente para proporcionar que a funcionalidade desejada e os benefícios de tais aplicações sejam atingidos. São exemplos de tais aplicações:

- Comércio eletrônico;
- Aplicações que afetam o faturamento e linhas de produção de empresas;
- Sistemas de transações Bancárias, entre outras aplicações.

Falhas podem causar impacto em redes que suportam esse tipo de aplicações, uma vez que podem afetar a operação da rede e causar perdas de produtividade, perdas de vidas, causar prejuízos e afetar custos, enfim, gerar situações não desejadas.

Por exemplo, a falha de um único enlace pode causar interrupções significativas no tráfego entre componentes de uma dada rede, principalmente, se a rede não provê caminhos alternativos para roteamento dinâmico. Tais interrupções podem gerar insatisfação aos clientes da rede que, por sua vez, requerem serviços de alto nível.

Vale ressaltar que alguns mecanismos podem ser empregados a fim de que os efeitos de falhas ou quedas, numa determinada rede de computadores, sejam reduzidos ou minimizados. Entre esses mecanismos, pode-se citar redundância como a estratégia base para qualquer outro mecanismo³ que aumente a disponibilidade (e também a confiabilidade) de redes de computadores. Tolerância a faltas [JALOTE94, SIEWIOREK-SWARZ92] tem por objetivo prover serviços apesar da presença de faltas (falhas são causadas devido a faltas) no sistema.

Redundância consiste em empregar recursos extras para manter operacional um sistema, redes de computadores, etc., apesar da ocorrência de falhas ou quedas. Redundância pode ser: física ou temporal [JALOTE94, SIEWIOREK-SWARZ92].

A redundância física refere-se a recursos de software e hardware extras. Um exemplo de redundância física é prover enlaces adicionais a uma rede de computadores,

² Partes de uma rede podem estar não operacionais devido a outras causas, tais como degradação de desempenho, congestionamento, entre outros fatores [KUBAT89, LI-SILVESTER84].

³ Os algoritmos de árvore de cobertura [PERLMAN92] e roteamento dinâmico [FORD AT AL.97, SOARES AT. AL.95] podem ser considerados mecanismos tradicionais utilizados para aumentar a disponibilidade de redes de computadores. Porém, há também soluções proprietárias, tais como HSRP (*Hot Standby Routing Protocol*) [CISCO96, CISCO99A] e DLD (*Deterministic Load Distribution*) [CISCO99B]. No entanto, apresentar detalhes a respeito de tais mecanismos não faz parte do escopo desta dissertação.

para permitir possível re-roteamento de uma mensagem, caso a recepção por seu destino, não tenha sido confirmada.

Na prática, componentes de uma rede podem falhar, servidores podem ficar sobrecarregados e causar impacto nos clientes, conexões podem ser perdidas e assim por diante. Além disso, o fato de existirem aplicações com requisitos específicos de disponibilidade gera a necessidade de projetar cuidadosamente redes para suportar tais aplicações.

Para tanto, métodos de avaliação de disponibilidade podem ser utilizados durante o processo de projeto de redes, a fim de que estratégias para aumentar a disponibilidade, topologias e tecnologias de redes sejam analisadas e escolhidas, de modo que a rede projetada possa atingir a disponibilidade requisitada pelas aplicações, considerando para tal, as metas de projeto da rede em questão.

Metas de projeto compreendem restrições do tipo minimizar custos e maximizar desempenho. Além disso, o projeto de redes deve considerar as necessidades operacionais de porções particulares da rede e os efeitos que falhas em componentes podem causar à disponibilidade total ou parcial da rede.

Por exemplo, em projetos de redes com restrições de custo, o projetista de rede pode utilizar um método para procurar topologias que garantam um certo nível de disponibilidade a um custo acessível, de acordo com a meta de custo para o projeto em questão. O custo de inserir redundância física, por exemplo, pode ser proibitivo. Sendo assim, projetistas devem analisar cuidadosamente a rede e definir quais porções requerem necessariamente redundância para que a disponibilidade especificada seja atingida.

A título de informação, uma das estratégias para projetar redes com requisitos específicos de disponibilidade é construir possíveis configurações topológicas da rede e comparar suas respectivas disponibilidades, em busca de atingir a disponibilidade pré-especificada para a rede em questão. Os artigos [*PIERRI-ELGIBAOU*97, *BALAKRISHNAN ET AL.*98] apresentam propostas de projeto de redes, neste cenário.

A seguir, apresenta-se uma discussão geral a respeito dos métodos de avaliação de disponibilidade de redes de computadores existentes.

2.2. Métodos de Avaliação de Disponibilidade de Redes de Computadores

Para descrever o escopo de abrangência e as suposições consideradas para avaliar a disponibilidade de uma determinada rede de computadores, os métodos de avaliação devem se basear num modelo específico.

Um modelo bastante utilizado é aquele no qual uma rede de computadores é representada por um grafo [SWAMY-THULASIRAMAN81], cujos vértices representam os dispositivos da rede e os arcos representam os enlaces de comunicação. Ao formular um modelo, as seguintes informações devem ser definidas:

- Se são consideradas falhas apenas em vértices, apenas em arcos ou em ambos;
- Quais os estados possíveis para os vértices e arcos. Métodos típicos de avaliação de disponibilidade consideram, basicamente, dois estados: operacional e não operacional (ou em falha). No entanto, pode-se considerar também o estado intermediário, ou seja, um vértice ou arco pode estar operacional, porém funcionando com um nível de desempenho degradado;
- O relacionamento entre as falhas, isto é, se são falhas dependentes ou independentes. A maioria dos métodos de avaliação de disponibilidade assumem que falhas são independentes. Entretanto, [AGGARWAL ET AL.75] é um exemplo de método de avaliação que considera falhas dependentes enquanto [KUBAT89] permite tratar falhas dependentes e independentes.

Um exemplo típico de modelo é dado a seguir. Uma rede de computadores é representada pelo grafo não direcionado $G = (V, A)$, composto por um conjunto finito V de N_V vértices e um conjunto finito A de N_A arcos. Cada arco consiste num par não ordenado de vértices, isto é, $a = (v_i, v_j)$ onde $v_i, v_j \in V$. Laços próprios não são permitidos, ou seja, não se considera arcos do tipo $a = (v_i, v_i)$. Os vértices e arcos estão em um dos estados: operacional ou não operacional. Estados intermediários não são considerados. O estado de um vértice ou arco é um evento aleatório, isto é, está sujeito a fenômenos aleatórios. O estado de cada vértice e de cada arco é independente do estado de qualquer outro vértice ou arco. A cada vértice v e a cada arco a está associada uma probabilidade de estar operacional $P(v)$ e $P(a)$, respectivamente.

Esse modelo de disponibilidade é particularmente considerado por [BALL-VAN

SLYKE77]. A maioria dos outros métodos de avaliação de disponibilidade que serão discutidos, neste capítulo, também seguem modelos similares ao apresentado acima.

Para avaliar a disponibilidade de uma rede de computadores, além de adotar um modelo específico, é necessário utilizar uma medida específica de disponibilidade. Uma medida tradicional de disponibilidade é conectividade. Diz-se que uma rede está operacional se ela está conectada, isto é, se qualquer dispositivo pode enviar mensagens para qualquer outro dispositivo ou para o resto da rede. Isto implica que os dispositivos e o caminho - uma seqüência alternada de dispositivos e enlaces - entre eles devem estar operacionais. A maioria dos métodos de avaliação de disponibilidade existentes baseiam-se em conectividade.

Tipicamente, as medidas de disponibilidade são as seguintes:

- Probabilidade da rede estar conectada. Uma rede está conectada quando todos os seus pares de dispositivos podem se comunicar [*AGGARWAL ET AL.75, BALL79, BALL-VAN SLYKE77, BAUER ET AL.87, KUBAT89, MYRVOLD96, VAN SLYKE-FRANK72, VAN SLYKE ET AL.75, TORRIERI94*];
- Probabilidade de todos os caminhos entre dois dispositivos específicos estarem operacionais [*BALL-VAN SLYKE77, BALL79, HÄNSLER ET AL.74*];
- Probabilidade de todos os dispositivos operacionais se comunicarem com um dispositivo específico [*BALL79, KERSHENBAUM-VAN SLYKE73*];
- Probabilidade de um número específico de dispositivos se comunicarem um com o outro, numa rede de computação distribuída [*KE-WANG97*];
- Número de pares de dispositivos que estão se comunicando [*BALL79, KE-WANG97, KUBAT89, VAN SLYKE ET AL.75*];
- Número de dispositivos que estão se comunicando com um dispositivo específico [*BALL79, KERSHENBAUM-VAN SLYKE73*];
- Fração de dispositivos se comunicando [*KUBAT89*].

Em termos de exatidão dos resultados gerados, os métodos de avaliação de disponibilidade podem ser classificados em métodos exatos e métodos heurísticos.

Métodos exatos de avaliação de disponibilidade, tais como [*HÄNSLER ET AL.74*,

MISRA-RAO70] podem ter complexidade computacional exponencial, com crescimento de acordo com o tamanho da rede em questão. Portanto, são métodos recomendados apenas para avaliar redes muito pequenas (poucas dezenas de dispositivos). Segundo [*VAN SLYKE ET AL.75*], de acordo com resultados experimentais, métodos exatos de avaliação da disponibilidade são indicados, geralmente, para avaliar redes com no máximo 30 componentes suscetíveis a falhas.

Para melhor esclarecer as afirmações anteriores, considere o seguinte exemplo: suponha que um método de avaliação de disponibilidade utilize uma técnica de análise de disponibilidade baseada na enumeração de todos os possíveis estados de falhas da rede. Vale salientar que um estado de falha da rede representa um conjunto de componentes que estão não operacionais. Os conjuntos {}, {1}, {2}, ..., {1, 2}, ..., {1, 2, 3} e assim por diante, onde {*i*} significa falha no componente *i*, são exemplos de estados de falha. Para uma rede com, por exemplo, 30 componentes seria necessário analisar 2^{30} ou 1.073.741.824 de estados. Para um número de componentes muito grande, tal análise seria impossível.

Entretanto, para redes estruturadas em árvore, há métodos de avaliação de disponibilidade baseados em técnicas recursivas que podem gerar resultados exatos e que possuem complexidade computacional polinomial. [*KERSHENBAUM-VAN SLYKE73*] é um exemplo desse tipo de método de avaliação.

Apesar de gerar resultados aproximados, métodos heurísticos não têm complexidade computacional exponencial. Ao contrário, normalmente têm complexidade computacional polinomial. Portanto, os métodos heurísticos são indicados para avaliar a disponibilidade de redes grandes. [*KUBAT89, LI-SILVESTER84, VAN SLYKE-FRANK72, VAN SLYKE ET AL.75*] são exemplos de métodos heurísticos de avaliação de disponibilidade de redes de computadores.

Em relação à complexidade computacional dos métodos tradicionais de avaliação da disponibilidade de redes, [*BALL80*] apresenta um estudo detalhado a respeito da dificuldade em computar algumas medidas tradicionais de disponibilidade.

Cada método de avaliação de disponibilidade utiliza técnicas particulares para a análise e cálculo da disponibilidade de redes. [*WILKOV72*] apresenta uma visão geral a respeito de alguns métodos tradicionais de avaliação de disponibilidade de redes, apresentando assim, aspectos referentes às medidas de disponibilidade utilizadas e as

técnicas básicas de análise e cálculo empregadas por tais métodos.

Alguns métodos de avaliação de disponibilidade, [VAN SLYKE-FRANK72, VAN SLYKE ET AL.75] por exemplo, baseiam-se em análise combinatória, ou seja, analisam combinações do tipo: número de formas que uma quantidade específica de enlaces com falha possam desconectar uma dada rede de computadores e, por fim, calculam a probabilidade de tal evento.

[HÄNSLER ET AL.74] utiliza a técnica de enumerar conjuntos de estados exclusivos (por exemplo, x estados contendo y componentes com falhas) e calcula a probabilidade de eventos relacionados.

[KUBAT89, BALL-VAN SLYKE77] baseiam-se na enumeração de todos os possíveis estados de falha, a fim de calcular a disponibilidade da rede enquanto [LI-SILVESTER84, BONAVENTURA80] considera apenas os estados mais prováveis da rede.

Em adição, outro ponto de consideração relevante, no contexto de avaliação de disponibilidade de redes de computadores, refere-se à consideração de ocorrência de falhas em componentes da rede. Alguns métodos de avaliação de disponibilidade, tais como [VAN SLYKE-FRANK72, HÄNSLER ET AL.74] consideram apenas falhas em enlaces; no entanto, sabe-se que tal suposição não é real. Dispositivos de redes falham e, sendo assim, os métodos de avaliação de disponibilidade devem considerar tal comportamento ao efetuar os cálculos. [VAN SLYKE ET AL.75, BALL79, KUBAT89, LI-SILVESTER84] são métodos de avaliação de disponibilidade que fazem tal consideração.

Por outro lado, pode-se derivar expressões de disponibilidade que consideram falhas em dispositivos e em enlaces a partir de expressões de disponibilidade geradas por métodos de avaliação de disponibilidade que consideram falhas apenas em enlaces. [AGGARWAL ET AL.75, TORRIERI94] propõem técnicas que permitem tal derivação.

Uma vez que redes de computadores atuais são constituídas por dispositivos e enlaces de alta disponibilidade, além de dispor enlaces alternativos para comunicação no evento de falhas, medidas de conectividade já não expressam completamente a real situação das redes.

Hoje em dia, uma rede falha não apenas quando ocorrem perdas de conectividade parcial ou total, mas também quando a rede atinge um nível de desempenho inaceitável. Por exemplo, a falha de certos componentes pode direta ou indiretamente, aumentar a

intensidade de tráfego, diminuir a velocidade de transmissão, causar atrasos inaceitáveis, causar qualidade inaceitável de transmissão, gerar congestionamento e perdas de informação, etc. [KUBAT86].

Por exemplo, em redes que suportam aplicações que requerem alta disponibilidade e em redes que integram voz, vídeo e dados, a degradação de desempenho pode causar impacto no custo e produtividade dos clientes tanto quanto falhas nos componentes da rede [3COM98].

[KUBAT89, LI-SILVESTER84] propõem métodos de avaliação de disponibilidade que consideram este novo cenário. Através de tais métodos, pode-se avaliar medidas de desempenho, tais como atraso e a própria disponibilidade, em redes de computadores constituídas por componentes suscetíveis a falhas. Ambos os métodos avaliam a disponibilidade a partir da análise dos estados de falha da rede. A medida de desempenho total é obtida a partir do valor de desempenho para cada estado considerado.

Dependendo da rede sendo analisada, pode-se escolher uma medida de disponibilidade, bem como um método de avaliação adequado para calculá-la, de modo que os requisitos reais de disponibilidade da rede em questão sejam atingidos e as considerações de falhas sejam adequadamente analisadas.

No próximo capítulo, uma nova medida de disponibilidade de redes de computadores baseadas na arquitetura cliente/servidor em n -camadas, é definida formalmente.

CAPÍTULO 3

UMA NOVA MEDIDA DE DISPONIBILIDADE DE REDES DE COMPUTADORES

Atualmente, redes corporativas são baseadas na arquitetura cliente/servidor em 3-camadas (ou n -camadas)¹ [EDWARDS99], isto é, redes corporativas são logicamente compostas por três módulos: camada de serviços dos usuários, camada intermediária e camada de dados.

A disponibilidade de serviços² de rede para os clientes depende de algumas restrições quanto à conectividade³ entre clientes e servidores e entre os próprios servidores, ou seja, quanto à conectividade entre as camadas, anteriormente citadas. Neste capítulo, apresenta-se uma nova medida de disponibilidade de redes de computadores baseada especificamente nessas conectividades parciais, ou seja, na disponibilidade da comunicação cliente/servidor em redes de computadores baseadas na arquitetura cliente/servidor em n -camadas.

¹ Vide apêndice B, para obter maiores informações a respeito da arquitetura cliente/servidor em n -camadas.

² Cada servidor de uma rede corporativa tem uma função específica. Então, diz-se que ele provê serviços específicos aos clientes da rede. Por exemplo, servidores de arquivos provêm serviços tais como armazenamento e compartilhamento de arquivos.

³ Diz-se que um dispositivo está conectado a outro dispositivo se eles estão conectados diretamente ou se houver um caminho (de dispositivos intermediários e enlaces) que os interconecte.

3.1. Avaliação de Disponibilidade de Redes de Computadores Baseadas na Arquitetura Cliente/Servidor em n -Camadas

Atualmente, redes corporativas têm muitos servidores e clientes, e baseiam-se no modelo clássico cliente/servidor⁴. Os servidores provêm serviços a todos ou a uma parte dos clientes da rede. Os clientes, por sua vez, têm acesso a, basicamente, dois tipos de servidores: servidores departamentais e servidores corporativos. Esses tipos de servidores provêm os principais serviços necessários às aplicações típicas de uma rede corporativa. Entre essas aplicações estão aplicações de correio eletrônico, compartilhamento de arquivos corporativos (compreendem os arquivos de interesse para todos ou para a maioria dos clientes da rede) ou departamentais (pertencentes a um grupo de trabalho específico), entre outras aplicações [3COM98].

Servidores departamentais suportam um grupo específico de clientes. Servidores de arquivos e de impressão são exemplos de servidores departamentais. Aplicações de colaboração entre grupos de trabalho, por exemplo, acessam servidores de arquivos a fim de compartilhar recursos, tais como arquivos. Esses servidores não são corporativos, já que não são acessados por todos os clientes da rede.

Servidores corporativos suportam todos ou a maioria dos clientes da rede. Servidores de correio eletrônico, servidores Web, servidores de bancos de dados corporativos são exemplos deste tipo de servidor.

No entanto, considerar uma rede corporativa como um conjunto de clientes e servidores é uma visão simplificada de uma rede de computadores corporativa. Hoje em dia, redes corporativas são estruturadas segundo a arquitetura cliente/servidor em 3-camadas (ou n -camadas). A figura 3.1 mostra um esquema que relaciona o modelo cliente/servidor em 2 camadas e sua implementação em n camadas.

⁴ No modelo cliente/servidor, denomina-se servidor a entidade que oferece serviços à entidade denominada cliente. Os servidores aceitam solicitações de serviços, executam a solicitação e enviam os resultados ao cliente solicitante. Por exemplo, um servidor de arquivos tem por função oferecer a seus clientes os serviços de armazenamento e acesso a informações, bem como compartilhamento de informações.

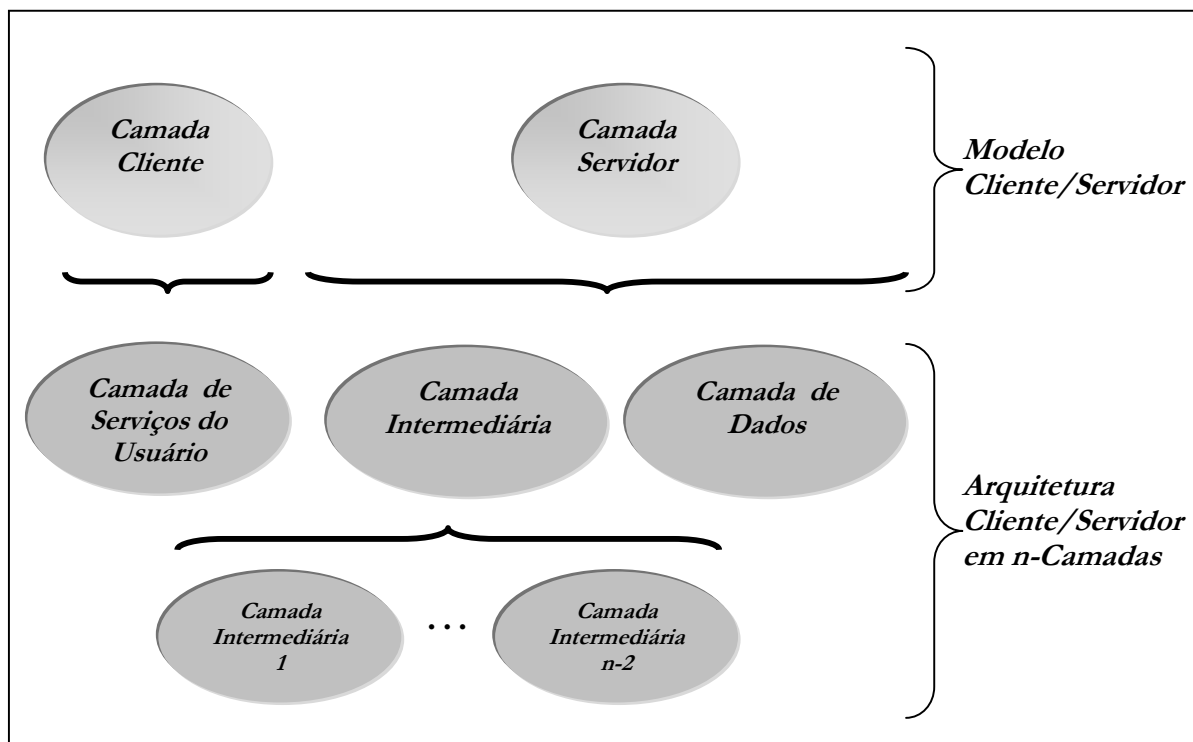


Figura 3.1 - Arquitetura Cliente/Servidor em n-Camadas

No modelo clássico cliente/servidor em 2 camadas, todo o processamento é efetuado no cliente que, por sua vez, acessa diretamente os servidores de bancos de dados que funcionam apenas como repositórios de dados. Sendo assim, é necessária uma conexão particular de cada cliente para o servidor de banco de dados. E, em adição, deve-se instalar em cada cliente, um software cliente específico, para permitir que os clientes acessem o servidor de banco de dados.

Na arquitetura cliente/servidor em 3 ou n -camadas, os clientes estão na camada de serviços do usuário e o lado servidor foi dividido em duas camadas: camada intermediária e camada de dados. Os clientes acessam, através de folheadores⁵, as aplicações localizadas na camada intermediária. Na camada de dados estão contidos, precisamente, os bancos de dados corporativos.

A camada intermediária pode ser implementada em várias camadas, especificamente, de 1 a $(n - 2)$ camadas. Uma das razões para utilizar a camada intermediária é compartilhar recursos caros e/ou escassos utilizados pelos clientes. Por

⁵ Software de visualização necessários para acessar as aplicações localizadas nos servidores corporativos da camada intermediária.

exemplo, já que não há mais a necessidade de uma conexão dedicada entre clientes e servidores de banco de dados, há economia computacional e de memória. Em adição, a instalação de softwares clientes específicos não é mais necessária, uma vez que a maioria do processamento é realizado na camada intermediária, sem necessidade de acessos diretos à camada de dados, por parte dos clientes.

Redes corporativas têm muitos clientes e, cada cliente, pode pertencer a um ou mais grupos de trabalho. E, para que estes clientes recebam os serviços oferecidos pelos servidores, as seguintes regras de comunicação devem ser seguidas: um cliente particular deve acessar os seus respectivos servidores departamentais e os servidores corporativos da camada intermediária e estes, por sua vez, devem acessar uns aos outros e os servidores corporativos da camada de dados. Estes últimos não precisam, tipicamente, acessar uns aos outros, precisam apenas acessar os servidores corporativos da camada intermediária.

Redes corporativas atuais são estruturadas segundo a arquitetura cliente/servidor em 3-camadas (ou n -camadas) e, portanto, possuem necessidades e restrições de disponibilidade que os métodos, discutidos no capítulo 2, não podem analisar completamente. Sendo assim, propõe-se uma nova medida de disponibilidade e um método de avaliação de disponibilidade específicos para avaliar disponibilidade de redes de computadores corporativas atuais.

Considerando as necessidades particulares de comunicação, descritas no capítulo 2, evidencia-se que a conectividade entre clientes e servidores (e entre os próprios servidores) é um fator crítico para a operação da rede, uma vez que os enlaces que conectam clientes e servidores (e os servidores entre si), os próprios servidores e equipamentos de interconexão podem falhar e afetar, de forma crítica, a disponibilidade de serviços para os clientes da rede.

O fato de não receber os serviços da rede pode causar séria insatisfação aos clientes, bem como à rede corporativa, como um todo. Por exemplo, para redes corporativas que suportam aplicações de comércio eletrônico, o fato dos serviços da rede não estarem disponíveis para um cliente em particular pode gerar perdas de negócios, afetando o lucro da empresa.

Uma vez que situações como estas podem frequentemente ocorrer em corporações, faz-se necessário medir a quantidade de clientes de uma determinada rede corporativa que

estão recebendo os serviços desejados. Em muitas situações, este é um fator relevante para o sucesso da empresa, já que o desempenho da empresa está diretamente relacionado ao desempenho da rede.

Falhas em um ou mais componentes podem afetar a conectividade da rede apenas parcialmente, isto é, tais falhas não implicam necessariamente que toda a rede vai ficar não operacional (como considerado ao definir medidas de disponibilidade baseadas em conectividade total). Ao invés disso, uma porção da rede pode estar operacional e recebendo serviços da rede, apesar de outra porção estar desconectada (ou seja, não operacional).

Redes corporativas típicas podem estar, muitas vezes, parcialmente conectadas e, no entanto, uma parte da rede pode continuar operacional. Sendo assim, medidas de disponibilidade baseadas em conectividade total não expressam situações do mundo real. Nesse caso, conectividades parciais devem ser consideradas.

Entretanto, embora algumas medidas tradicionais de disponibilidade sejam baseadas em conectividade parcial entre componentes, nenhuma delas expressa exatamente as restrições de comunicação requeridas para o contexto de redes corporativas baseadas na arquitetura cliente/servidor em n -camadas (descritas detalhadamente na seção 3.2).

Muitas vezes, falhas podem afetar a disponibilidade de serviços para um determinado cliente, sem afetar os demais clientes. Sendo assim, a nova medida de disponibilidade expressa uma fração de clientes que estão recebendo serviços de uma rede de computadores baseada na arquitetura cliente/servidor em 3-camadas (ou n -camadas).

3.2. Escopo Considerado

Hoje em dia, redes corporativas suportam aplicações tais como comércio eletrônico e sistemas de transações bancárias, entre outras aplicações, que dependem de uma infraestrutura de alta disponibilidade para que os resultados obtidos sejam satisfatórios tanto para os clientes quanto para a empresa, como um todo.

Para que uma rede de computadores possa prover uma infra-estrutura de alta disponibilidade é necessário que, entre outros pontos, os clientes estejam conectados a seus servidores. Na verdade, a disponibilidade dos serviços da rede para os clientes depende da disponibilidade dos próprios servidores e dos caminhos que os interconectam. Por exemplo,

aplicações cooperativas entre grupos de trabalho dependem da disponibilidade dos servidores de arquivos, nos quais os arquivos compartilhados estão localizados, assim como da disponibilidade dos caminhos entre os clientes e os servidores da rede. Sendo assim, os clientes constituintes dos grupos de trabalho em questão devem estar conectados aos servidores, a fim de que as tarefas de cooperação entre os grupos sejam efetuadas, de modo satisfatório.

No entanto, atualmente, redes corporativas baseiam-se na arquitetura cliente/servidor em 3-camadas (ou n -camadas) para estruturar "logicamente" a comunicação entre clientes e servidores corporativos da rede, ou seja, a rede está estruturada segundo as seguintes camadas: camada de serviços dos usuários (clientes), camada intermediária (servidores corporativos, tais como servidores Web, servidores de aplicação executando *middleware* e servidores de correio eletrônico) e camada de dados (servidores de bancos de dados). Sendo assim, para disponibilizar serviços aos clientes de uma rede nesse cenário, algumas restrições, quanto às conectividades entre clientes e servidores (e entre servidores) são impostas.

Um cliente particular está recebendo serviços completos da rede se as cinco condições, descritas a seguir, são satisfeitas:

- O cliente precisa acessar os seus próprios servidores departamentais;
- O cliente precisa acessar os servidores corporativos da camada intermediária;
- Os servidores corporativos da camada intermediária precisam acessar uns aos outros. Por exemplo, o servidor Web pode disparar a execução de uma aplicação no servidor de aplicações;
- Os servidores corporativos da camada intermediária precisam acessar os servidores corporativos da camada de dados;
- Os servidores corporativos da camada de dados precisam acessar os servidores corporativos da camada intermediária.

Entretanto, os componentes de uma rede típica falham com alguma probabilidade. Sendo assim, as conectividades parciais, anteriormente descritas, necessárias para disponibilizar serviços de rede aos clientes são afetadas por falhas nos componentes da rede.

Neste trabalho, considera-se que os seguintes componentes de rede são suscetíveis a falhas:

- Hospedeiros (clientes, servidores departamentais, servidores corporativos da camada intermediária e servidores corporativos da camada de dados);
- Equipamentos de interconexão (concentradores, comutadores, roteadores, etc.);
- Enlaces individuais.

Falhas em hospedeiros, equipamentos de interconexão ou enlaces podem causar impacto nas conectividades parciais necessárias para disponibilizar serviços de rede aos clientes. Para entender melhor tais efeitos, considere a rede exemplo da figura 3.2.

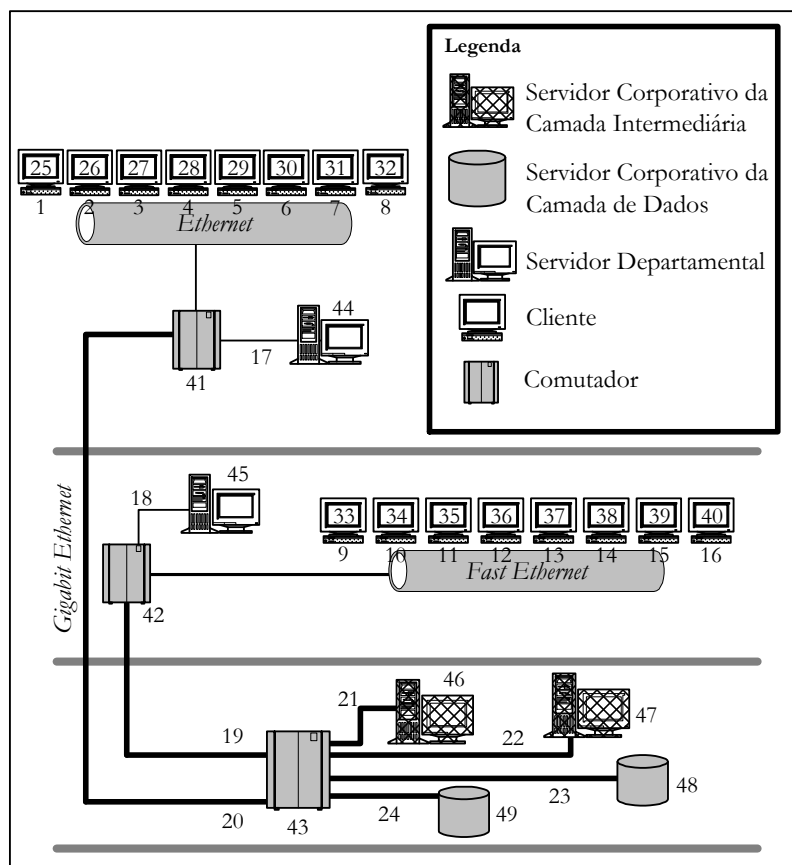


Figura 3.2 - Rede de Computadores Exemplo

A rede da figura 3.2 é hipotética e possui 49 componentes. Particularmente, têm-se 16 clientes, 2 servidores departamentais, 2 servidores corporativos da camada intermediária e 2 servidores corporativos da camada de dados. Os componentes são identificados por uma

numeração ordenada de acordo com a seqüência: enlaces, clientes, equipamentos de interconexão, servidores departamentais, servidores corporativos da camada intermediária e servidores corporativos da camada de dados.

Para os clientes da rede-exemplo da figura 3.2, qualquer falha nos servidores departamentais (servidor 44 para os clientes da sub-rede *Ethernet* - 25 a 32 - e servidor 45 para os clientes da sub-rede *Fast Ethernet* - 33 a 40), nos servidores corporativos (servidores 46, 47, 48 e 49) ou nos caminhos (incluindo os enlaces individuais e os equipamentos de interconexão) que conectam clientes e servidores pode causar impacto na disponibilidade de serviços aos clientes da rede. Algumas falhas em componentes, no exemplo da figura 3.2, são críticas. Por exemplo, uma falha no comutador 43 afeta a comunicação entre os servidores corporativos e todos os clientes da rede e, portanto, estes não poderão receber os serviços completos da rede.

Determinadas falhas podem afetar diretamente as conectividades parciais necessárias para disponibilizar serviços de rede apenas para um cliente, entretanto, os outros clientes não serão necessariamente afetados. Considere a rede exemplo da figura 3.2. Falhas do enlace 17, do comutador 41 ou do servidor departamental 44 afetam diretamente os clientes da sub-rede *Ethernet*, não causando impacto aos clientes da sub-rede *Fast Ethernet*. Da mesma maneira, falhas do enlace 18, do comutador 42 ou do servidor departamental 45 afetam apenas os clientes da sub-rede *Fast Ethernet*, não causando efeitos aos clientes da sub-rede *Ethernet*.

Diante desse cenário, a nova medida expressa especificamente a fração de clientes que estão recebendo serviços completos da rede, ou seja, o número de clientes que estão recebendo serviços da rede em relação ao número total de clientes da rede. A seguir, apresenta-se a definição formal da nova medida de disponibilidade de redes de computadores baseadas na arquitetura cliente/servidor em n -camadas.

3.3. Definição Formal da Nova Medida de Disponibilidade

Utiliza-se a seguinte tabela de notações, a fim de definir formalmente a nova medida de disponibilidade de redes de computadores.

Notação

n_t - número total de componentes
 n_c - número de clientes
 n_{sd} - número de servidores departamentais
 n_{sci} - número de servidores corporativos da camada intermediária
 n_{scd} - número de servidores corporativos da camada de dados
 n_e - número de enlaces individuais
 n_{eq} - número de equipamentos de interconexão
 n_h - número total de hospedeiros ($n_c + n_{sd} + n_{sci} + n_{scd}$)
 n_{ef} - número de estados de falha da rede
 c - cliente
 sd - servidor departamental
 sci - servidor corporativo da camada intermediária
 scd - servidor corporativo da camada de dados
 e - enlace individual
 eq - equipamento de interconexão
 SD_i - conjunto dos servidores departamentais do cliente i
 S - conjunto de todos os servidores departamentais
 MAO - Matriz de Adjacências Originais
 MAE - Matriz de Adjacências de um Estado particular
 MCD - Matriz de Conectividades Desejadas
 MCE - Matriz de Conectividades reais de um Estado particular

Para considerar as conectividades parciais entre os clientes e servidores (e entre servidores) de uma rede de computadores baseada na arquitetura cliente/servidor em n -camadas, utilizam-se as seguintes matrizes: matriz de adjacências originais, matriz de conectividades desejadas, matriz de adjacências de um estado particular e matriz de conectividades reais de um estado particular.

Para representar as adjacências originais entre hospedeiros e equipamentos de interconexão de uma rede de computadores particular utiliza-se uma matriz de adjacências originais.

Considere que $MAO = [mao_{ij}]$ seja a matriz de adjacências originais de ordem $(n_h + n_{eq}) \times (n_h + n_{eq})$, onde $1 \leq i, j \leq (n_h + n_{eq})$. Essa matriz tem a estrutura apresentada na figura 3.3.

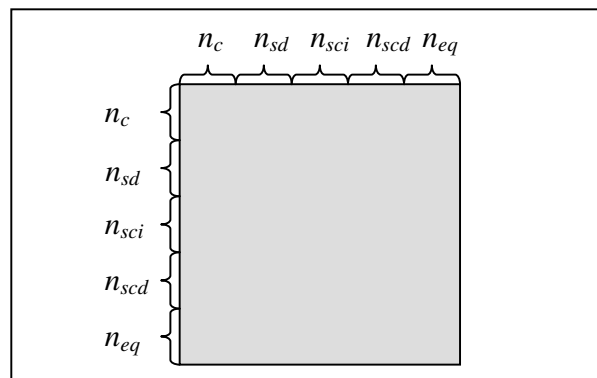


Figura 3.3 - Estrutura da Matriz de Adjacências Originais

Onde,

$$mao_{ij} = \begin{cases} 1, & \text{se o dispositivo } i \text{ é adjacente ao dispositivo } j \\ 0, & \text{caso contrário.} \end{cases}$$

Observe que os hospedeiros e equipamentos de interconexão são numerados na matriz, segundo a ordem: clientes, servidores departamentais, servidores corporativos da camada intermediária, servidores corporativos da camada de dados e equipamentos de interconexão.

A matriz de adjacências originais para a rede-exemplo da figura 3.2 tem os valores mostrados na tabela 3.1.

	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0

40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
41	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0
42	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	1	0	1	0	0	0	0
43	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	1	1	1
44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
46	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0

Tabela 3.1 - Matriz de Adjacências Originais da Rede-Exemplo 3.2

A matriz de conectividades desejadas é especificamente empregada para representar as conectividades parciais necessárias para que os serviços completos sejam disponibilizados aos clientes da rede.

No entanto, cada cliente tem necessidades próprias quanto às conectividades necessárias entre os hospedeiros, para que possa receber os serviços completos da rede. Por exemplo, todos os clientes da rede da figura 3.2 precisam estar conectados aos servidores corporativos 46, 47, 48 e 49, para que possam receber os serviços da rede. Entretanto, os clientes da rede *Ethernet* precisam estar conectados também ao servidor departamental 44 e os clientes da rede *Fast Ethernet* precisam estar conectados ao servidor departamental 45. Supondo que o servidor 44 é o servidor departamental dos clientes da rede *Ethernet* e o servidor 45 é o servidor departamental de rede *Fast Ethernet*. Além disso, para que um cliente particular esteja recebendo serviços completos da rede, não interessam as conectividades entre os outros clientes e o resto da rede. Portanto, utiliza-se uma matriz de conectividades desejadas para cada cliente da rede.

Portanto, a matriz de conectividades desejadas representa as conectividades necessárias entre os hospedeiros da rede, para que um cliente particular possa receber completamente os serviços da rede. Uma matriz de conectividades desejadas tem a estrutura mostrada na figura 3.4.

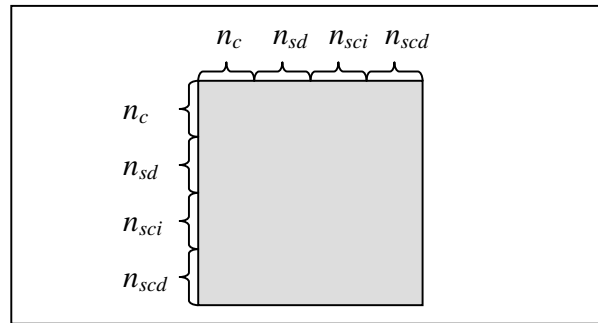


Figura 3.4 - Estrutura da Matriz de Conectividades Desejadas

Os hospedeiros estão numerados, na matriz, de acordo com a seguinte ordem: clientes, servidores departamentais, servidores corporativos da camada intermediária e servidores corporativos da camada de dados.

Seja $MCD^i = [mcd_{jk}^i]$ uma matriz de ordem $n_h \times n_h$, onde $1 \leq i \leq n_c$, $1 \leq j, k \leq n_h$, a matriz de conectividades desejadas para o i -ésimo cliente. O valor de mcd_{jk}^i pode ser 0 ou 1, dependendo da necessidade de conectividade entre os hospedeiros, identificados por j e k . O fato de mcd_{jk}^i ter valor 0, significa que os hospedeiros j e k não precisam estar conectados para que o cliente i esteja recebendo os serviços da rede, ou seja, se houver falhas de conexão entre os dois hospedeiros (ou se estes falharem), a disponibilidade de serviços para o cliente i não sofrerá impacto.

No entanto, caso mcd_{jk}^i tenha valor 1, qualquer falha de conexão entre os hospedeiros j e k (ou nestes) causa impacto na disponibilidade de serviços para o cliente i . Além disso, para o cliente particular i , não interessa se os outros clientes estão recebendo serviços da rede.

Portanto, os valores de mcd_{jk}^i necessários para que o cliente i esteja recebendo os serviços da rede são mostrados na tabela 3.2.

mcd_{jk}^i	Restrições	Significado
mcd_{kj}^i	$j > k, 1 \leq j \leq n_h, 1 \leq k \leq n_h$	A conectividade nos dois sentidos é necessária
1	$j = k = i$	O cliente i está conectado a si mesmo
0	$1 \leq j \leq n_c, j < k \leq n_h, j \neq i$	A conectividade dos outros clientes não

		interessa ao cliente i
0	$j = i, j < k \leq n_c$	O cliente i não precisa estar conectado a outros clientes
1	$j = i, n_c + 1 \leq k \leq n_c + n_{sd}, S_k \in SD_i$	O cliente i está conectado a todos os seus servidores departamentais, onde $SD_i = \{j \mid S_j \text{ é servidor departamental do cliente } i\}$ e S é o conjunto de todos os servidores departamentais da rede (S_j é o j -ésimo servidor departamental)
1	$j = i, n_c + n_{sd} + 1 \leq k \leq n_c + n_{sd} + n_{sci}$	O cliente i está conectado a todos os servidores corporativos da camada intermediária
0	$j = i, n_c + n_{sd} + n_{sci} + 1 \leq k \leq n_h$	O cliente i não precisa estar conectado aos servidores corporativos da camada de dados
0	$n_c + 1 \leq j \leq n_c + n_{sd}, j < k \leq n_h$	Os servidores departamentais não precisam estar conectados entre si ou estar conectados aos servidores corporativos da camada intermediária ou da camada de dados
1	$n_c + n_{sd} + 1 \leq j \leq n_c + n_{sd} + n_{sci}, j < k \leq n_h$	Os servidores corporativos da camada intermediária estão conectados entre si e conectados com os servidores corporativos da camada de dados
0	$n_c + n_{sd} + n_{sci} + 1 \leq j \leq n_h, j < k \leq n_h$	Os servidores corporativos da camada de dados não precisam estar conectados entre si

Tabela 3.2 - Definição do Valor de mcd_{jk}^i

A matriz de conectividades desejadas para o cliente 25 da rede-exemplo da figura 3.2 tem os valores mostrados na tabela 3.3.

	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	44	45	46	47	48	49
25	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0

26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
46	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1
47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1
48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0
49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0

Tabela 3.3 - Matriz de Conectividades Desejadas para o Cliente 25 da Rede-Exemplo 3.2

Para calcular a fração de clientes que estão recebendo serviços da rede, deve-se comparar as conectividades desejadas e as conectividades reais entre os hospedeiros da rede, a fim de determinar se os serviços da rede estão disponíveis para cada um dos clientes da rede.

As conectividades reais entre os hospedeiros da rede estão sujeitas aos efeitos das falhas nos próprios hospedeiros, nos equipamentos de interconexão e nos enlaces individuais da rede e, por consequência, nos caminhos que conectam todos os componentes de uma rede de computadores particular. Para levar em conta tais ocorrências de falhas, consideram-se estados de falhas da rede.

Estados de falha são combinações de componentes com falhas, ou seja, um estado de falha particular é constituído por um ou mais componentes com falha (isto é, no estado não operacional). Por exemplo, um estado de falha para a rede-exemplo da figura 3.2 poderia ser {20, 44, 48}, ou seja, nesse estado de falha o enlace 20, o servidor departamental 44 e o servidor corporativo da camada de dados 48 estão não operacionais.

Utiliza-se uma matriz de adjacências para cada estado de falha, para representar as adjacências entre hospedeiros e equipamentos que representam os efeitos de falhas em todos os componentes do estado de falha em questão.

Seja $MAE^{es} = [mae^{es}_{ij}]$ a matriz de adjacências para um estado de falha e tem a mesma estrutura que a matriz de adjacências originais, mostrada na figura 3.3, e é de ordem $(n_h + n_{eq}) \times (n_h + n_{eq})$, onde $1 \leq es \leq n_{ef}$, $1 \leq i, j \leq (n_h + n_{eq})$. Essa matriz representa as adjacências entre os hospedeiros e equipamentos de interconexão da rede, ao considerar o efeito dos componentes com falhas contidos no estado de falha e .

Portanto,

$$mae_{ij} = \begin{cases} 1, & \text{se o dispositivo } i \text{ é adjacente ao dispositivo } j, \\ 0, & \text{caso contrário.} \end{cases}$$

Por exemplo, suponha que um dos estados de falha para a rede da figura 3.2 seja $E = \{19, 25, 41, 46\}$. Sendo assim, a matriz de adjacências para o estado de falha E tem os valores mostrados na tabela 3.4.

	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

conectividades reais para um estado particular é obtida através da verificação do efeito das falhas nos componentes, presentes no estado de falha específico, que afetam as conectividades entre os hospedeiros da rede.

Seja $MCE^{es} = [mce^{es}_{ij}]$, onde $1 \leq es \leq n_{ef}$, $1 \leq i, j \leq n_h$. MCE^{es} é de ordem $n_h \times n_h$ e tem uma estrutura semelhante à estrutura da matriz de conectividades desejadas, apresentada na figura 3.3. Os valores de mce^{es}_{ij} são obtidos segundo a regra abaixo:

$$mce^{es}_{ij} = \begin{cases} 1, & \text{se o hospedeiro } i \text{ está conectado ao hospedeiro } j, \\ 0, & \text{caso contrário.} \end{cases}$$

Considerando ainda o estado de falha $E = \{19, 25, 41, 46\}$ para a rede da figura 3.2, a matriz de conectividades real para E contém os seguintes valores:

	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	44	45	46	47	48	49
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	1	0	0	0	0
34	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	1	0	0	0	0
35	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	1	0	0	0	0
36	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	1	0	0	0	0
37	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	1	0	0	0	0
38	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	1	0	0	0	0
39	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	1	0	0	0	0

40	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	1	0	0	0	0
44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
45	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	1	0	0	0	0
46	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Tabela 3.5 - Matriz de Conectividades Reais para o Estado de falha E da rede 3.2

As matrizes anteriormente descritas são necessárias para armazenar as informações sobre as conectividades entre os componentes da rede e são também utilizadas para definir formalmente a nova medida de disponibilidade de redes de computadores.

Para definir formalmente a nova medida de disponibilidade de redes de computadores baseadas na arquitetura cliente/servidor em n -camadas, consideram-se os possíveis estados de falhas da rede. Em uma rede com n_t componentes suscetíveis a falhas, há 2^{n_t} possíveis estados de falha. Os estados de falha da rede são denotados por E_k , onde $1 \leq k \leq 2^{n_t}$. E_k representa o k -ésimo estado da rede, onde $E_1 = \phi$, $E_2 = \{1\}$, $E_3 = \{2\}$ e assim por diante. Sendo assim, E_1 corresponde a um estado sem falhas enquanto $E_{2^{n_t}}$ corresponde ao estado onde todos os componentes da rede estão com falha.

A cada estado de falha E_k da rede, associa-se uma probabilidade $P(E_k)$ [LIL- SILVESTER84]. Tal probabilidade é calculada a partir da seguinte equação:

$$P(E_k) = \prod_{i=1}^{n_t} p_i (q_i/p_i)^{T_i(E_k)}$$

Onde p_i ⁶ corresponde à probabilidade do componente i ($1 \leq i \leq n_t$) estar operacional e $q_i = 1 - p_i$. Na equação:

⁶ Vale ressaltar que este valor não representa a probabilidade do componente estar num estado de falha particular, ao invés disso, p_i é probabilidade do componente estar operacional. Por exemplo, comutadores típicos têm disponibilidades entre 99.9 e 99.999%, ou seja, probabilidades de estar operacionais entre 0.999 e 0.99999.

$$T_i(E_k) = \begin{cases} 0, & \text{se o componente } i \text{ está operacional no estado } E_k \\ 1, & \text{caso contrário.} \end{cases}$$

Especificamente, a nova medida de disponibilidade de redes de computadores, ou seja, a fração de clientes que estão recebendo serviços de rede é definida por:

$$D = \sum_{k=1}^{2^{n_t}} P(E_k) D(E_k)$$

Onde, $D(E_k)$ é a disponibilidade de todos os clientes, considerando o estado de falha E_k , e é definida por:

$$D(E_k) = [\sum_{i=1}^{n_c} D_i(E_k)] / n_c$$

$D_i(E_k)$ corresponde à disponibilidade de serviços completos da rede para o cliente específico i e seu valor é determinado a partir da seguinte equação:

$$D_i(E_k) = \begin{cases} 1, & \text{se } MCD^i \text{ AND } MCE^{E_k} = MCD^i \\ 0, & \text{caso contrário.} \end{cases}$$

AND é uma operação booleana lógica⁷. Tal operação é realizada para comparar, célula por célula, todos os valores da matriz de conectividades desejadas para um cliente específico MCD^i com a matriz de conectividades reais para o estado de falha E_k - MCE^{E_k} . Portanto, $MCD^i \text{ AND } MCE^{E_k} = MCD^i$, se e somente se, $mcd_{jt}^i \text{ AND } mce^{E_k} = mcd_{jt}^i \forall 1 \leq j, t \leq n_h$. Uma vez que essa restrição é atendida, o cliente i tem suas conectividades desejadas satisfeitas e, conseqüentemente, está recebendo os serviços completos da rede, considerando o estado de falha E_k .

Considere a matriz de conectividades desejadas para o cliente 25 (tabela 3.3.) e a matriz de conectividades reais para o estado $E = \{19, 25, 41, 46\}$ (tabela 3.5) da rede 3.2. Já que o cliente 25 pertence ao estado de falha E , a comparação entre a primeira célula da matriz de conectividades desejadas e a primeira célula da matriz de conectividades reais

⁷ $0 \text{ AND } 0 = 0$
 $1 \text{ AND } 0 = 0$
 $0 \text{ AND } 1 = 0$
 $1 \text{ AND } 1 = 1$

não atende à restrição ($mcd_{jt}^i \text{ AND } mce^{Ek} = mcd_{jt}^i \forall 1 \leq j, t \leq n_h$), uma vez que $mcd_{jt}^i = 1$ e $mce^{Ek} = 0$. Portanto, isto significa que o cliente 25 não está recebendo serviços completos da rede, ao considerar-se o estado de falha E .

Vale ressaltar que, considerou-se apenas a arquitetura cliente/servidor em 3-camadas (camada de serviços de rede, camada intermediária e camada de dados) ao definir a nova medida de disponibilidade. No entanto, a camada intermediária da arquitetura cliente/servidor em n -camadas pode ser dividida "logicamente" em sub-camadas, precisamente de 1 a $n-2$ sub-camadas. Sendo assim, pode-se afirmar que a nova medida de disponibilidade de redes é genérica para o contexto de redes de computadores baseadas na arquitetura cliente/servidor em n -camadas⁸, devido a dois motivos:

- 1) As conectividades parciais necessárias para disponibilizar serviços de rede aos clientes não são modificadas ao considerar uma rede com um número maior de camadas intermediárias, uma vez que a divisão da camada intermediária é uma divisão lógica e à nova medida interessa necessariamente à conectividade com os servidores corporativos, além das outras conectividades (seção 3.2) independente do número de camadas em que estão inseridos;
- 2) A matriz de conectividades desejadas pode ser utilizada para representar as conectividades necessárias à disponibilidade de serviços em redes baseadas na arquitetura cliente/servidor para qualquer valor de n , uma vez que não restringe especificamente a um valor específico, o número de camadas intermediárias consideradas.

Avaliar a disponibilidade de uma rede de computadores considerando os 2^{n_t} possíveis estados de falha impõe limites ao tamanho da rede avaliada. Métodos de avaliação de disponibilidade baseados na enumeração de todos os possíveis estados de falhas, como mencionado no capítulo 2, são recomendados para redes de pequeno porte (poucas dezenas de componentes).

⁸ Esta afirmação vale para o modelo cliente/servidor tradicional, isto é, em 2-camadas. Neste caso, a matriz de conectividades desejadas, para cada cliente de uma determinada rede de computadores, seria constituída por clientes e servidores.

Sendo assim, no próximo capítulo, propõe-se um método heurístico de avaliação de disponibilidade de redes de computadores que utiliza a nova medida de disponibilidade proposta considerando, no entanto, apenas os estados de falhas mais prováveis da rede.

CAPÍTULO 4

UM MÉTODO DE AVALIAÇÃO DE DISPONIBILIDADE DE REDES DE COMPUTADORES BASEADAS NA ARQUITETURA CLIENTE/SERVIDOR EM N-CAMADAS

No capítulo anterior, foi definida uma nova medida de disponibilidade de redes de computadores baseadas na arquitetura cliente/servidor em n -camadas. Para avaliar a disponibilidade, utilizando tal medida, consideram-se todos os possíveis estados de falha da rede, segundo a sua definição formal - seção 3.3.

No entanto, métodos de avaliação de disponibilidade de redes de computadores que se baseiam na enumeração de todos os possíveis estados de falha são recomendados apenas para redes pequenas, por terem complexidade computacional exponencial.

Sendo assim, propõe-se um método heurístico de avaliação de disponibilidade de redes, que utiliza a nova medida de disponibilidade proposta, e considera apenas os estados de falha mais prováveis da rede, obtendo assim, um valor aproximado para a medida de disponibilidade. E com isso, o método de avaliação de disponibilidade tem complexidade computacional polinomial.

Este capítulo está organizado de acordo com a seguinte seqüência: define-se o modelo considerado ao definir o método de avaliação; em seguida, apresenta-se formalmente o método de avaliação de disponibilidade; e finalmente, determina-se a sua complexidade computacional.

4.1. Modelo Considerado

O método de avaliação de disponibilidade de redes de computadores baseadas na arquitetura cliente/servidor em n -camadas baseia-se no seguinte modelo: uma rede de

computadores é representada por um grafo¹ não direcionado² $G = (V, A)$ contendo n_v vértices e n_a arcos. $G = (V, A)$ compreende um conjunto finito de vértices $V = \{v_1, v_2, \dots, v_{n_v}\}$ e um conjunto finito de arcos $A = \{a_1, a_2, \dots, a_{n_a}\}$, onde cada arco é identificado por um par de vértices, ou seja, $a_k = (v_i, v_j)$, onde $1 \leq k \leq n_a$ e $1 \leq i, j \leq n_v$. Laços próprios³ são permitidos.

Os vértices representam os clientes, servidores e equipamentos de interconexão enquanto os arcos representam os enlaces individuais das redes de computadores. Os vértices e arcos podem estar em um dos estados: operacional ou em falha. Estados intermediários, tais como estado operacional com nível degradado de desempenho, não são considerados.

A cada componente (vértice ou arco) i está associada uma probabilidade p_i (de estar no estado operacional) e uma probabilidade $q_i = 1 - p_i$ (de estar no estado em falha).

Os estados são considerados eventos aleatórios, ou seja, os componentes da rede estão sujeitos a influências ou a mecanismos de mudança aleatórios [BLAKE79]. Além disso, o estado de um vértice ou arco é totalmente independente do estado de qualquer outro vértice ou arco.

A seguir, apresentam-se algumas considerações necessárias à definição formal do método de avaliação de disponibilidade de redes de computadores baseadas na arquitetura cliente/servidor em n -camadas.

4.2. Considerações Preliminares para a Definição Formal do Método de Avaliação de Disponibilidade

A tabela de notações abaixo é utilizada ao definir o método de avaliação de disponibilidade de redes de computadores.

Notação

n_t	- número total de componentes ($n_h + n_e + n_{eq}$)
n_{ef}	- número de estados mais prováveis da rede
n_c	- número de estações clientes
n_{sd}	- número de servidores departamentais
n_{sci}	- número de servidores corporativos da camada intermediária
n_{scd}	- número de servidores corporativos da camada de dados
n_e	- número de enlaces individuais

¹ Para obter informações detalhadas a respeito da teoria dos grafos, consulte [SWAMY-THULASIRAMAN81].

² Um grafo não direcionado contém arcos identificados por pares não ordenados de vértices.

³ Permitir laços próprios significa permitir que os vértices sejam adjacentes a eles próprios, isto é, $a_k = (v_i, v_i)$, onde $1 \leq k \leq n_a$ e $1 \leq i \leq n_v$.

n_{eq} - número de equipamentos de interconexão
 n_h - número de hospedeiros ($n_c + n_{sd} + n_{sci} + n_{scd}$)
 c - cliente
 sd - servidor departamental
 sci - servidor corporativo da camada intermediária
 scd - servidor corporativo da camada de dados
 eq - equipamento de interconexão
 e - enlace individual
 E - estado de falha particular
 MAO - matriz de adjacências originais
 MAE - matriz de adjacências de um estado particular
 MCD - matriz de conectividades desejadas
 MCE - matriz de conectividades reais de um estado particular
 CSD - estrutura cliente_servidor_departamental
 TC - estrutura transitive_closure

Considere o contexto, a seguir. Seja $G = (V, A)$ o grafo que representa uma rede de computadores baseada na arquitetura cliente/servidor em n -camadas composta por n_t componentes. Onde,

$$V = \{c_1, \dots, c_{n_c}, sd_1, \dots, sd_{n_{sd}}, sci_1, \dots, sci_{n_{sci}}, scd_1, \dots, scd_{n_{scd}}, eq_1, \dots, eq_{n_{eq}}\}$$

$$A = \{e_1, \dots, e_{n_e}\}$$

Além disso, assume-se que:

1. O componente i , onde $1 \leq i \leq n_t$, está no estado operacional com probabilidade p_i e no estado em falha com probabilidade $q_i = 1 - p_i$;
2. Os componentes estão ordenados segundo $R_1 \geq R_2 \geq \dots \geq R_{n_t}$, onde $R_i = q_i/p_i$.

A nova medida de disponibilidade, ou seja, a fração de clientes que estão recebendo serviços completos da rede, é obtida através da seguinte equação:

$$D = \sum_{k=1}^{2^{n_t}} P(E_k) D(E_k) \quad (\text{Equação 1})$$

Onde $P(E_k)$ é a probabilidade (de estar operacional) associada ao estado de falha E_k , onde $1 \leq k \leq 2^{n_t}$. Tal probabilidade é calculada a partir da seguinte equação:

$$P(E_k) = \prod_{i=1}^{n_t} p_i (q_i/p_i)^{T_i(E_k)} \quad (\text{Equação 2})$$

$$T_i(E_k) = \begin{cases} 0, & \text{se o componente } i \text{ está operacional no estado de falha } E_k \\ 1, & \text{caso contrário.} \end{cases}$$

E,

$$D(E_k) = [\sum_{i=1}^{n_c} D_i(E_k)]/n_c$$

Observe que, segundo a definição formal da nova medida de disponibilidade, seu valor é obtido considerando-se todos os possíveis estados de falha da rede, ou seja, 2^{n_t} estados de falha. Os estados de falha, portanto, são denotados por $E_1 = \phi$, $E_2 = \{1\}$, $E_3 = \{2\}$ e assim por diante.

Sendo assim, E_1 corresponde a um estado sem falhas enquanto $E_2^{n_t}$ corresponde ao estado onde todos os componentes estão no estado em falha.

Porém, métodos de avaliação de disponibilidade de redes de computadores, baseados na enumeração de todos os possíveis estados de falha, são recomendados apenas para redes de pequeno porte (poucas dezenas de componentes), uma vez que para uma rede de computadores com n_t componentes suscetíveis a falhas, há 2^{n_t} possíveis estados de falha e, por conseqüência, tais métodos de avaliação de disponibilidade têm complexidade computacional exponencial.

Sendo assim, neste capítulo, propõe-se um método heurístico de avaliação de disponibilidade de redes de computadores baseadas na arquitetura cliente/servidor em n -camadas que, ao invés de considerar todos os possíveis estados de falha, considera apenas um subconjunto do espaço amostral de estados de falha.

Para tanto, tal método de avaliação baseia-se na proposta de enumeração dos estados de falha mais prováveis da rede apresentada por *Li* e *Silvester* em [*LI-SILVESTER83*, *LI-SILVESTER84*].

Para determinar os estados de falha mais prováveis da rede, comparam-se as probabilidades associadas aos estados de falha da rede, à medida que esses são gerados. Tais estados de falha são gerados de acordo com uma ordem decrescente de probabilidades.

Considere os n_{ef} estados de falha mais prováveis da rede, denotados por E_k , onde $1 \leq k \leq n_{ef}$. Os n_{ef} estados de falha mais prováveis são gerados, segundo a ordem $E_1, E_2, \dots, E_{n_{ef}}$ tal que $P(E_1) \geq P(E_2) \geq \dots \geq P(E_{n_{ef}})$. Caso contrário, seria necessário varrer todo o espaço amostral de estados de falha, em busca do n_{ef} estados de falha mais prováveis da rede.

O valor de n_{ef} , ou seja, o número de estados de falha mais prováveis da rede pode ser um valor pré-fixado ou um valor suficiente para cobrir uma determinada porção do espaço amostral de estados de falha da rede em questão. Por exemplo, para uma rede de computadores com n_t componentes, pode-se considerar $n_{ef} = n_t$, $n_{ef} = 10n_t$, $n_{ef} = 100n_t$ etc.

A seguir, descreve-se o algoritmo de geração dos n_{ef} estados de falha mais prováveis de uma rede de computadores - algoritmo ORDER - proposto por Li e Silvester em [LI-SILVESTER83, LI-SILVESTER84].

Seja $A = \{E_x, E_y, \dots, E_z\}$ um conjunto ordenado de estados de falha tal que $P(E_x) \geq P(E_y) \geq \dots \geq P(E_z)$, onde E_i corresponde ao estado de falha i , onde $1 \leq i \leq n_{ef}$ e $P(E_i)$ é sua probabilidade associada. Os conjuntos B_j e A_k , onde $1 \leq j, k \leq n_t$, são conjuntos de estados de falha similares ao conjunto de estados de falha A .

Algoritmo 1 - Geração de Estados de Falha mais Prováveis de uma Rede de Computadores

Objetivo: Gerar os n_{ef} estados de falha mais prováveis da rede.

Entrada: conjuntos de estados de falha A, A_{n_t}, B_{n_t} e o valor de n_{ef} .

Saída: o conjunto A_{n_t} contendo os n_{ef} estados de falha mais prováveis da rede.

ORDER

{

i' = primeiro inteiro, tal que $2^{i'} \geq n_{ef}$

$A_0 = \phi$;

$B_0 = \phi$;

$A = \phi$;

Primeira fase do algoritmo

De $i = 1$ até i' faça

{

$B_{i-1} = \mathbf{anexa}(A_{i-1}, B_{i-1}, i)$;

$A_i = \mathbf{insere}(A_{i-1}, B_{i-1})$;

}

O número de estados de falha gerados no final do estágio i é 2^i

Segunda fase do algoritmo

De $i = i' + 1$ até n_t faça

{

$B_{i-1} = \mathbf{anexa}(A_{i-1}, B_{i-1}, i)$;

$A = \mathbf{insere}(A_{i-1}, B_{i-1})$;

$A_i = \mathbf{seleciona}(A, n_{ef})$;

```

    }
    #  $A_{n_i}$  contém os  $n_{ef}$  estados de falha mais prováveis da rede #
}
anexa (A, B, i)
{
    B = A  $\cup$  {i}
    = { $E_x \cup \{i\}$ ,  $E_y \cup \{i\}$ , ...,  $E_z \cup \{i\}$ }
    # A probabilidade de  $P(E_j \cup \{i\}) = P(E_j) * R_i$ ,  $\forall E_j \in A$ . Sendo assim, o conjunto de estados de
falha B também está ordenado #
}
insere (A, B)
{
    C = {conjunto ordenado de estados de falha obtido com a inserção de cada
elemento do conjunto ordenado de estados de falha B no conjunto ordenado de estados
de falha A}
}
seleciona (A,  $n_{ef}$ )
{
    C = {conjunto ordenado que contém os  $n_{ef}$  estados de falha mais prováveis da
rede, obtido a partir do conjunto ordenado de estados de falha A}
    # Observe que  $C \subset A$  #
}

```

Pode-se definir limites superior e inferior para a nova medida de disponibilidade, uma vez que os estados de falha são gerados em ordem decrescente de probabilidade e, isto permite uma convergência rápida entre os limites, à medida que estados de falha adicionais são considerados. Portanto, tem-se um valor aproximado da nova medida de disponibilidade sem ter de considerar todos os possíveis estados de falha da rede.

Tal definição de limites superior e inferior baseia-se no modelo genérico de limites superior e inferior para uma medida de desempenho, proposto por *Li e Silvester*⁴ em [*LI-SILVESTER84*], que se fundamenta na seguinte proposição: seja M , uma medida de desempenho qualquer, e seja $M(E_k)$ o seu valor quando a rede está no estado de falha

⁴ *Li e Silvester* basearam-se em [*VAN SLYKE-FRANK72*] ao propor limites superior e inferior para uma medida de desempenho de rede de computadores.

E_k . Segundo [LI-SILVESTER84], a relação entre os limites superior e inferior de M é dada por:

$$M(E_2^{n_t}) \leq M(E_k) \leq M(E_1), \text{ onde } 1 \leq k \leq 2^{n_t} \quad (\text{Equação 3})$$

Portanto, o desempenho da rede é melhor quando se considera o estado de falha E_1 , isto é, quando se consideram todos os componentes no estado operacional. No entanto, o desempenho da rede é pior quando se considera o estado $E_2^{n_t}$, uma vez que todos os componentes estão em falha, nesse estado.

Os valores dos limites superior e inferior de uma medida de desempenho qualquer da rede, ao considerar os n_{ef} estados de falha mais prováveis da rede, são obtidos a partir das equações:

$$M_{inferior} = \sum_{k=1}^{n_{ef}} P(E_k)M(E_k) + (1 - \sum_{k=1}^{n_{ef}} P(E_k))M(E_2^{n_t})$$

$$M_{superior} = \sum_{k=1}^{n_{ef}} P(E_k)M(E_k) + (1 - \sum_{k=1}^{n_{ef}} P(E_k))M(E_1)$$

Para o caso específico da nova medida de disponibilidade, têm-se os seguintes limites superior e inferior

$$D_{inferior} = \sum_{k=1}^{n_{ef}} P(E_k)D(E_k) + (1 - \sum_{k=1}^{n_{ef}} P(E_k))D(E_2^{n_t})$$

$$D_{superior} = \sum_{k=1}^{n_{ef}} P(E_k)D(E_k) + (1 - \sum_{k=1}^{n_{ef}} P(E_k))D(E_1)$$

Uma vez que $E_2^{n_t}$ é o estado de falha constituído por todos os componentes suscetíveis a falhas da rede, pode-se afirmar que $D(E_2^{n_t}) = 0$. E_1 é um estado sem falhas, então $D(E_1) = 1$. Portanto,

$$D_{inferior} = \sum_{k=1}^{n_{ef}} P(E_k)D(E_k)$$

$$D_{superior} = \sum_{k=1}^{n_{ef}} P(E_k)D(E_k) + (1 - \sum_{k=1}^{n_{ef}} P(E_k))$$

Para calcular tais limites superior e inferior da nova medida de disponibilidade de redes de computadores baseadas na arquitetura cliente/servidor em n -camadas, define-se formalmente um método de avaliação de disponibilidade, apresentado a seguir.

4.3. Definição Formal do Método de Avaliação de Disponibilidade de Redes Baseadas na Arquitetura Cliente/Servidor em n -Camadas

Inicialmente, apresentam-se os tipos de dados e estruturas utilizadas para definir formalmente o método de avaliação de disponibilidade.

Tipos de Dados e Estruturas

n_{ef} é o número de estados de falha mais prováveis da rede

INTEIRO *n_{ef}* ;

$P[n_i]$ é o vetor contendo as probabilidades dos componentes da rede estarem operacionais, ou seja, o componente i tem probabilidade $P[i]$ de estar operacional

$disponibilidade_total_clientes$ corresponde à fração de clientes que estão recebendo serviços da rede, ao considerar-se um estado de falha específico

PONTO FLUTUANTE *$P[n_i]$, $disponibilidade_total_clientes$* ;

$disponibilidade_cliente$ é VERDADEIRO, se o cliente sob consideração está recebendo serviços da rede e FALSO, caso contrário

BOOLEANO *$disponibilidade_cliente$* ;

$tipo_dispositivo$ é um tipo específico utilizado para identificar os hospedeiros e equipamentos de interconexão da rede

$lista_hospedeiros_equipamentos$ contém os identificadores dos clientes, servidores departamentais, servidores corporativos da camada intermediária, servidores corporativos da camada de dados e equipamentos de interconexão, nesta ordem

$lista_hospedeiros$ contém os identificadores dos clientes, servidores departamentais, servidores corporativos da camada intermediária e servidores corporativos da camada de dados, nesta ordem

TIPO_DISPOSITIVO *$lista_hospedeiros_equipamentos[n_h + n_{eq}]$, $lista_hospedeiros[n_h]$* ;

Armazenando os identificadores dos clientes de cada servidor departamental da rede

ESTRUTURA *$cliente_servidor_departamental$*

{

INTEIRO *$número_clientes_do_servidor_departamental$* ;

$tipo_clientes$ é um tipo específico utilizado para identificar os clientes da rede

TIPO_CLIENTES *$clientes[número_clientes_do_servidor_departamental]$* ;

} **CSD** *$[n_{sd}]$* ;

Armazenando os identificadores dos clientes, servidores departamentais, servidores corporativos da camada intermediária, servidores corporativos da camada de dados e equipamentos de interconexão, nesta ordem, e os valores das adjacências entre tais dispositivos, além de armazenar os identificadores dos enlaces que os conectam

ESTRUTURA *$matriz_adjacências$*

```

{
    # origem corresponde ao identificador do dispositivo origem #
    TIPO_DISPOSITIVO origem;
    # destino corresponde ao identificador do dispositivo destino #
    TIPO_DISPOSITIVO destino;
    # O valor de adjacência é 1, se o dispositivo origem é adjacente ao dispositivo destino #
    BOOLEANO adjacência;
    # Caso adjacência tenha valor 1, número_enlaces_adjacência contém o número de enlaces que
    conectam os dispositivos origem e destino #
    INTEIRO número_enlaces_adjacência;
    # tipo_enlace é um tipo específico utilizado para identificar os enlaces individuais da rede #
    TIPO_ENLACE enlaces[número_enlaces_adjacência];
}MAO[nh + neq][nh + neq];

# Armazenando os valores da matriz de adjacências para um estado de falha particular da rede #
ESTRUTURA matriz_adjacências_estado
{
    ESTRUTURA matriz_adjacências matriz_adjacências_estado[nh + neq][nh + neq];
}MAE[nef];

# Armazenando as conectividades específicas entre um hospedeiro origem e um hospedeiro destino da
rede #
ESTRUTURA conectividade
{
    # origem corresponde ao identificador do dispositivo origem #
    TIPO_DISPOSITIVO origem;
    # destino corresponde ao identificador do dispositivo destino #
    TIPO_DISPOSITIVO destino;
    # conectividade é 1, se o dispositivo origem está conectado ao dispositivo destino #
    BOOLEANO conectividade;
}

# Armazenando os valores das conectividades entre os hospedeiros da rede #
ESTRUTURA matriz_conectividades
{
    # Nas matrizes de conectividades (seja a matriz de conectividades desejadas, seja a matriz de
    conectividades para um estado de falha particular da rede), os identificados dos hospedeiros são

```

armazenados segundo a ordem: clientes, servidores departamentais, servidores corporativos da camada intermediária e servidores corporativos da camada de dados #

ESTRUTURA *conectividade* *matriz_conectividades*[n_h][n_h];

}*MCD*[n_c], *MCE*[n_{ef}];

Armazenando os componentes em falha que constituem um estado de falha particular, além de sua probabilidade de falha associada

ESTRUTURA *elementos*

{

INTEIRO *número_componentes_em_falha*;

tipo_componentes é um tipo específico para identificar os componentes da rede, ou seja, os hospedeiros, equipamentos de interconexão e enlaces individuais da rede

TIPO_COMPONENTES *componentes_em_falha*[*número_componentes_em_falha*];

PONTO FLUTUANTE *probabilidade_estado*;

}

Armazenando um conjunto de estados de falha, conjuntos de estados, tais como $A[n_i]$, $B[n_i]$ e A . Esses conjuntos são utilizados durante a geração de estados de falha mais prováveis da rede - algoritmo ORDER

ESTRUTURA *estados*

{

estados é um vetor de estados de falha, cujo tamanho não é pre-fixado

ESTRUTURA *elementos* *estados*[];

}*A*[n_i], *B*[n_i], *A*;

Armazenando os valores das adjacências entre dispositivos da rede, numa estrutura específica a ser utilizada durante a execução do algoritmo transitive closure - descrito a seguir

ESTRUTURA *transitive_closure*

{

origem corresponde ao identificador do dispositivo origem

TIPO_DISPOSITIVO *origem*;

destino corresponde ao identificador do dispositivo destino

TIPO_DISPOSITIVO *destino*;

O valor de adjacência é 1, se o dispositivo origem é adjacente ao dispositivo destino

BOOLEANO *adjacência*;

}*TC*[$n_h + n_{eq}$][$n_h + n_{eq}$];

Armazenando os valores dos limites superior e inferior da nova medida de disponibilidade de redes de computadores, considerando cada estado da rede. Armazenando, além disso, a probabilidade de cada estado de falha considerado - *probabilidade_acumulada_estados*

ESTRUTURA *limites_estados*

```
{  
    # limite_superior contém o valor do limite superior da nova medida de disponibilidade para um  
    estado de falha particular #  
    PONTO FLUTUANTE limite_superior;  
    # limite_inferior contém o valor do limite inferior da nova medida de disponibilidade para um  
    estado de falha particular #  
    PONTO FLUTUANTE limite_inferior;  
    PONTO FLUTUANTE probabilidade_acumulada_estados;  
}limites;
```

A seguir, descreve-se formalmente o método de avaliação de disponibilidade de redes de computadores baseadas na arquitetura cliente/servidor em n -camadas.

Algoritmo 2 - Método de Avaliação da Disponibilidade de Redes de Computadores Baseadas na Arquitetura Cliente/Servidor em n -Camadas

Objetivo: calcular os limites superior e inferior da nova medida de disponibilidade de redes de computadores baseadas na arquitetura cliente/servidor em n -camadas, ou seja, fração de clientes que estão recendo os serviços da rede.

Entrada: número de clientes, número de servidores departamentais, número de servidores corporativos da camada intermediária, número de servidores corporativos da camada de dados, número total de componentes, tipo de cada componente e sua respectiva probabilidade de estar operacional (esta lista deve estar em ordem crescente de probabilidades), lista de servidores departamentais e seus respectivos clientes, lista de adjacências entre os hospedeiros e equipamentos de interconexão.

Saída: os valores dos limites superior e inferior da nova medida de disponibilidade de redes, ao considerar os n_{ef} estados de falha mais prováveis da rede.

MÉTODO DE AVALIAÇÃO DE DISPONIBILIDADE

```

{
  # Obtendo a lista de clientes de cada um dos servidores departamentais da rede #
  obtem_clientes_dos_servidores_departamentais_da_rede(CSD[ $n_{sd}$ ],  $n_{sd}$ );

  # Obtendo a matriz de adjacências originais da rede #
  obtem_matriz_adjacencias_originais_da_rede(MAO[ $n_h + n_{eq}$ ][ $n_h + n_{eq}$ ],  $n_h$ ,  $n_{eq}$ );

  # Obtendo a matriz de conectividades desejadas para cada cliente da rede #
  obtem_matriz_conectividades_desejadas_dos_clientes_da_rede(MCD[ $n_c$ ][ $n_h$ ][ $n_h$ ]
,  $n_c$ ,  $n_h$ );

  # Considerando o primeiro estado de falha da rede, isto é,  $E_1$  #
   $E_1 = \phi$ ;
   $A_0 = \{E_1\}$ ;

  # Calculando a probabilidade do estado de falha  $E_1$  #
  calcula_probabilidade_do_primeiro_estado_de_falha_da_rede(A[0]);

  # Gerando os  $n_{ef}$  estados de falha mais prováveis da rede - Algoritmo 1 - previamente descrito #
  executa_algoritmo_ORDER;

  # Inicializando a estrutura que armazena os limites superior e inferior da nova medida de
  disponibilidade de redes #
  limites.limite_inferior = 0;
  limites.limite_superior = 0;
  limites.probabilidade_acumulada_estados=0;

  # Considerando os  $n_{ef}$  estados de falha mais prováveis da rede #
  De  $i = 1$  até  $n_{ef}$  faça
  {
    # Gerando a matriz de adjacência para o  $i$ -ésimo estado de falha da rede #
    gera_matriz_adjacencias_para_i-esimo_estado_de_falha_da_rede(MAE[ $n_{ef}$ ],
 $n_h$ ,  $n_{eq}$ );

    # Copiando a matriz de adjacência do  $i$ -ésimo estado de falha da rede para a estrutura TC #

```

```

copia_matriz_adjacências_i-ésimo_estado_de_falha_para_TC(TC[nh + neq][
nh + neq], nh, neq, MAE[i]);

# Executando o algoritmo Transitive Closure5 #
executa_algoritmo_transitive_closure(TC[nh + neq][nh + neq], nh, neq);

# Gerando a matriz de conectividades reais para o i-ésimo estado de falha da rede #
gera_matriz_conectividades_i-ésimo_estado_de_falha_da_rede(MCE[nef], nh,
neq);

# Calculando os limites superior e inferior da nova medida de disponibilidade de rede #
calcula_limites_superior_e_inferior_da_nova_medida_disponibilidade(MCD
[nc], MAE[nef], disponibilidade_total_clientes, nc, nh);

# Calculando os valores do limite inferior da nova medida de disponibilidade #
disponibilidade_total_clientes = disponibilidade_total_clientes / nc;
limites.limite_inferior = limites.limite_inferior +
A[nt].estados[i].probabilidade_estado *
disponibilidade_total_clientes);
limites.probabilidade_acumulada_estados =
limites.probabilidade_acumulada_estados +
A[nt].estados[i].probabilidade_estado;
}

# Após considerar todos os nef estados de falha mais prováveis da rede, limites.limite_inferior
contém o valor do limite inferior da nova medida de disponibilidade de redes. De modo similar,
limites.limite_superior contém o valor do limite superior da nova medida de disponibilidade #
limites.limite_superior = limites.limite_inferior + (1 -
limites.probabilidade_acumulada_estados);
}

```

⁵ Dada a matriz de adjacência A do grafo $G = (V, E)$, o algoritmo *transitive closure* gera a matriz de adjacência A^+ do *transitive closure* de G . O *transitive closure* de G é o grafo $G^+ = (V, E^+)$, onde um arco (i, j) está em E^+ , se e somente se há um caminho de i para j , ou seja, há uma seqüência de vértices v_0, \dots, v_t com $t > 0$, $i = v_0$ e $(v_r, v_{r+1}) \in E^+$ para todo $r < t$ e $v_t = j$ [Even79].

obtem_clientes_dos_servidores_departamentais_da_rede(CSD[n_{sd}])

```
{  
  De  $i = 1$  até  $n_{sd}$  faça  
  {  
    CSD[ $i$ ].numero_clientes_do_servidor_departamental =  
obtem_numero_clientes_do_servidor_departamental();  
    CSD[ $i$ ].clientes = obtem_clientes_do_servidor_departamental();  
  }  
}
```

obtem_matriz_adjacencias_originais_da_rede(MAO[$n_h + n_{eq}$][$n_h + n_{eq}$], n_h , n_{eq})

```
{  
  De  $i = 1$  até ( $n_h + n_{eq}$ ) faça  
  {  
    De  $j = 1$  até ( $n_h + n_{eq}$ ) faça  
    {  
      mao[ $i$ ][ $j$ ].origem = lista_hospedeiros_equipamentos [ $i$ ];  
      mao[ $i$ ][ $j$ ].destino = lista_hospedeiros_equipamentos [ $j$ ];  
      mao[ $i$ ][ $j$ ].adjacencia = obtem_adjacencia_original();  
      mao[ $i$ ][ $j$ ].numero_enlaces_adjacencia = obtem_numero_enlaces_adjacencia();  
      mao[ $i$ ][ $j$ ].enlaces = obtem_enlaces_adjacencia();  
    }  
  }  
}
```

**obtem_matriz_conectividades_desejadas_dos_clientes_da_rede(MCD[n_c][n_h][n_h],
 n_c , n_h)**

```
{  
  De  $i = 1$  até  $n_c$  faça  
  {  
    De  $j = 1$  até  $n_h$  faça  
    {  
      De  $k = 1$  até  $n_h$  faça  
      {
```

```

        mcd[i].matriz_conectividades[j][k].origem = lista_hospedeiros[j];
        mcd[i].matriz_conectividades[j][k].destino = lista_hospedeiros[k];
        mcd[i].matriz_conectividades[j][k].conectividade =
obtem_conectividade_desejada();
    }
}
}
}

```

calcula_probabilidade_do_primeiro_estado_de_falha_da_rede(A[0])

```

{
    De i = 1 até  $n_t$  faça
    {
        A[0].estados[0].probabilidade_estado = A[0].estados[0].probabilidade_estado *
P[i];
    }
}

```

gera_matriz_adjacencias_para_i-ésimo_estado_de_falha_da_rede(MAE[n_{ef}], n_h , n_{eq})

```

{
    De j = 1 até ( $n_h + n_{eq}$ ) faça
    {
        De k = 1 até ( $n_h + n_{eq}$ ) faça
        {
            mae[i].matriz_adjacencias_estado[j][k].origem =
lista_hospedeiros_equipamentos[j];
            mae[i].matriz_adjacencias_estado[j][k].destino =
lista_hospedeiros_equipamentos[k];
            mae[i].matriz_adjacencias_estado[j][k].adjacencia =
obtem_adjacencias_estado();
            mae[i].matriz_adjacencias_estado[j][k].numero_enlaces =
obtem_numero_enlaces();

```

```

    mae[i].matriz_adjacências_estado[j][k].enlaces =
obtem_enlaces_adjacência();
    }
}
}

copia_matriz_adjacências_i-ésimo_estado_de_falha_para_TC(TC[nh + neq][ nh +
neq], nh, neq, MAE[i])
{
    De j=1 até (nh + neq) faça
    {
        De k = 1 até (nh + neq) faça
        {
            TC[j][k].origem = lista_hospedeiros_equipamentos[j];
            TC[j][k].destino = lista_hospedeiros_equipamentos[k];
            TC[j][k].adjacência = mae[i].matriz_adjacências_estado[j][k].adjacência;
        }
    }
}

executa_algoritmo_transitive_closure(TC[nh + neq][nh + neq], nh, neq)
{
    De j=1 até (nh + neq) faça
    {
        De k = 1 até (nh + neq) faça
        {
            De t = 1 até (nh + neq) faça
            {
                TC[k][t].adjacência = (TC[k][t].adjacência) OR ( TC[k][j].adjacência
                AND TC[j][t].adjacência);
            }
        }
    }
}

```

}

gera_matriz_conectividades_i-ésimo_estado_de_falha_da_rede($MCE[n_{ef}]$, $TC[n_h + n_{eq}]$ [$n_h + n_{eq}$], n_h , n_{eq})

{

De $j=1$ até $(n_h + n_{eq})$ faça

{

De $k=1$ até $(n_h + n_{eq})$ faça

{

$mce[i].matriz_conectividades[j][k].origem = lista_hospedeiros[j];$

$mce[i].matriz_conectividades[j][k].destino = lista_hospedeiros[k];$

$mce[i].matriz_conectividades[j][k].conectividade = TC[j][k].adjacência;$

}

}

}

calcula_limites_superior_e_inferior_da_nova_medida_disponibilidade($MCD[n_c]$, $MAE[n_{ef}]$, n_c , n_h , $disponibilidade_total_clientes$)

{

$disponibilidade_total_clientes = 0;$

De $j=1$ até n_c faça

{

De $k=1$ até n_h faça

{

De $t=1$ até n_h faça

{

Se ($(mae[i].matriz_conectividades[k][t].conectividade$ **E**

$mcd[j].matriz_conectividades[k][t].conectividade)$ **for igual a**

$mcd[j].matriz_conectividades[k][t].conectividade)$) **faça**

{

$disponibilidade_cliente = VERDADEIRO;$

}

caso contrário

}

```

    {
        disponibilidade_cliente = FALSO;
        sai do laço que varre as matrizes;
    }
}
}
Se disponibilidade_cliente for igual a VERDADEIRO faça
{
    disponibilidade_total_clientes = disponibilidade_total_clientes + 1;
}
}
}

```

A seguir, determina-se a complexidade computacional do método de avaliação da disponibilidade de redes de computadores baseadas na arquitetura cliente/servidor em n -camadas, apresentado anteriormente.

4.4. Complexidade Computacional

A complexidade computacional de um algoritmo refere-se às necessidades computacionais do algoritmo, tais como número de operações lógicas ou aritméticas requeridas, para solucionar um determinado problema.

Para determinar a complexidade computacional do algoritmo 2, proposto na seção anterior, analisam-se os seus módulos constituintes. A seguir, tais módulos são identificados e sua complexidade computacional particular é apresentada.

<i>Módulo</i>	<i>Complexidade Computacional</i>
Obtenção da lista de clientes de cada um dos servidores departamentais da rede	$O(n_{sd})$
Obtenção da matriz de adjacências originais da rede	$O[(n_h + n_{eq})^2]$
Obtenção da matriz de conectividades desejadas para cada cliente da rede	$O(n_c n_h^2)$
Cálculo da probabilidade do primeiro estado de falha da rede	$O(n_i)$

Geração dos estados de falha mais prováveis da rede [LI-SILVESTER83, LI-SILVESTER84]	$O(n_t^2 n_{ef} + n_t n_{ef} \log n_{ef})$
Geração das matrizes de adjacências dos estados mais prováveis da rede	$O[n_{ef}(n_h + n_{eq})^2]$
Cópia das matrizes de adjacências dos estados de falha mais prováveis da rede	$O[n_{ef}(n_h + n_{eq})^2]$
Execução do algoritmo <i>transitive closure</i>	$O[n_{ef}(n_h + n_{eq})^3]$
Geração das matrizes de conectividades reais para os estados de falha mais prováveis da rede	$O[n_{ef}(n_h + n_{eq})^2]$
Cálculo dos limites superior e inferior da nova medida de disponibilidade	$O(n_{ef} n_c n_h^2)$

Tabela 4.1 - Complexidade Computacional das várias etapas do Método de Avaliação de Disponibilidade de Redes de Computadores Baseadas na Arquitetura Cliente/Servidor em n-Camadas

Portanto,

$$O[n_{sd} + (n_h + n_{eq})^2 + n_c n_h^2 + n_t + n_{ef} n_t^2 + n_t n_{ef} \log n_{ef} + n_{ef}(n_h + n_{eq})^2 + n_{ef}(n_h + n_{eq})^2 + n_{ef}(n_h + n_{eq})^3 + n_{ef}(n_h + n_{eq})^2 + n_{ef} n_c n_h^2] \text{ ou,}$$

$$O[n_{sd} + (n_h + n_{eq})^2 + n_c n_h^2 (1 + n_{ef}) + n_t + n_{ef} n_t^2 + n_t n_{ef} \log n_{ef} + 3 n_{ef}(n_h + n_{eq})^2 + n_{ef}(n_h + n_{eq})^3]$$

Uma vez que o valor de n_t é maior que os valores de n_{sd} , n_c , n_h , n_{eq} , já que $n_t = n_h + n_{eq} + n_e$, então pode-se utilizar a seguinte expressão:

$$O[2n_t + 2n_t^2 + n_t^3 + n_t^3 n_{ef} + n_t + n_{ef} n_t^2 + n_{ef} n_t \log n_{ef} + 3 n_{ef}(2n_t)^2 + n_{ef}(2n_t)^3] \text{ ou,}$$

$$O[2n_t + 2n_t^2 + n_t^3 + n_t^3 n_{ef} + n_{ef} n_t^2 + n_{ef} n_t \log n_{ef} + 3 n_{ef}(2n_t)^2 + n_{ef}(2n_t)^3]$$

Utilizando as propriedades particulares da notação $O(\cdot)$ [CAMPELO-MACULAN94, KRONSSJO87, SMITH89], pode-se afirmar que a complexidade computacional do método de avaliação de disponibilidade proposto é:

$$O[2n_t + 2n_t^2 + n_t^3 + n_{ef} n_t^3 + n_{ef} n_t^2 + n_{ef} n_t \log n_{ef} + 3 n_{ef}(2n_t)^2 + n_{ef}(2n_t)^3]$$

$$O[n_{ef} n_t^3 + n_{ef} n_t \log n_{ef}]$$

Portanto, $O[n_{ef} n_t^3 + n_{ef} n_t \log n_{ef}]$, isto é, um algoritmo de complexidade computacional polinomial.

No próximo capítulo, são apresentados os resultados experimentais obtidos ao empregar o método de avaliação de disponibilidade de redes de computadores, proposto neste capítulo, no contexto de algumas configurações topológicas particulares.

CAPÍTULO 5

ESTUDOS DE CASO

Métodos de avaliação de disponibilidade podem ser utilizados em processos de projetos de redes de computadores, a fim de obter estimativas de disponibilidade para as redes analisadas considerando, para tanto, o efeito das possíveis falhas em seus componentes.

Neste capítulo, emprega-se o método heurístico de avaliação de disponibilidade, definido no capítulo anterior, para avaliar a disponibilidade de algumas configurações topológicas particulares considerando um número específico de estados de falhas mais prováveis da rede. Com isso, caracteriza-se o comportamento da disponibilidade total da rede utilizando a nova medida de disponibilidade de redes, ou seja, a fração de clientes que estão recebendo serviços completos numa rede de computadores baseada na arquitetura cliente/servidor em n -camadas.

O método de avaliação de disponibilidade de redes foi implementado na linguagem de programação *C* [KERNIGHAN-RITCHIE88] e as execuções foram realizadas numa máquina com as seguintes características: *Pentium* II, 300 MHz, 192 M de memória RAM, 390 M de área de paginação (*swap*), disco rígido IDE de 8G e SCSI de 4G, executando GNU/Linux Red Hat 6.2.

Este capítulo está organizado na seguinte seqüência: inicialmente, apresenta-se um estudo a respeito do comportamento dos limites superior e inferior da nova medida de disponibilidade em relação ao número de estados de falhas mais prováveis considerados. Em seguida, o impacto da adição de clientes, servidores departamentais e corporativos a uma rede de computadores é estudado. Apresenta-se também um estudo a respeito do efeito da variação das disponibilidades de todos os componentes e dos enlaces individuais para os limites superior e inferior da nova medida de disponibilidade. Estuda-se o impacto da adição de redundância física (redundância de enlaces e de equipamentos de interconexão, particularmente) e em seguida, apresenta-se um estudo a respeito dos

impactos causados à disponibilidade da rede, à medida que há uma mudança na estrutura topológica de uma determinada rede, ou seja, quando uma rede *collapsed backbone* passa a ser estruturada hierarquicamente. Por fim, apresentam-se os resultados experimentais obtidos em relação ao tempo de processamento do método de avaliação de disponibilidade, ao considerar o tamanho da rede e o número de estados de falha mais prováveis considerados, para configurações topológicas particulares.

A seguir, descrevem-se as considerações utilizadas ao realizar os estudos de caso.

5.1. Considerações para os Estudos de Caso

As considerações aqui apresentadas representam o escopo considerado para realizar os estudos de caso. As redes de computadores baseiam-se na arquitetura cliente/servidor em n -camadas e possuem as seguintes características:

- Os componentes considerados suscetíveis a falhas são: clientes, servidores departamentais, servidores corporativos da camada intermediária e da camada de dados, equipamentos de interconexão (concentradores, comutadores e roteadores) e enlaces individuais;
- O tamanho das configurações topológicas estudadas varia de 40 a 145 componentes;
- As disponibilidades atribuídas aos componentes variam na faixa de 0.999 a 0.9999. Vale ressaltar que os componentes típicos de uma rede de computadores possuem, atualmente, disponibilidades que variam em torno de 0.9995 [*3COM99*, *3COM00*, *CISCO99A*, *CISCO99B*, *CISCO99C*, *CISCO00*].

Em grande parte das configurações topológicas utilizadas nos estudos caso, os componentes da rede possuem as disponibilidades listadas na tabela 5.1.

<i>Componente</i>	<i>Disponibilidade (%)</i>
Enlaces Individuais	99.95
Clientes	99.95
Equipamentos de Interconexão	99.98
Servidores Departamentais	99.98
Servidores Corporativos da Camada Intermediária	99.98
Servidores Corporativos da Camada de Dados	99.98

Tabela 5.1 - Disponibilidades dos Componentes das Configurações Topológicas Utilizadas nos Estudos de Caso

Ao variar as disponibilidades dos componentes da rede (seção 5.4), a tabela anterior não será seguida. Além disso, são atribuídas disponibilidades aos componentes da rede, de acordo com a seguinte ordem:

disponibilidades dos enlaces individuais \leq disponibilidades dos clientes \leq disponibilidades dos equipamentos de interconexão \leq disponibilidades dos servidores departamentais \leq disponibilidades dos servidores corporativos da camada intermediária \leq disponibilidades dos servidores corporativos da camada de dados

Vale ressaltar que os componentes de uma rede podem ter qualquer valor de disponibilidade, isto é, os valores das disponibilidades dos clientes, servidores, equipamentos de interconexão e enlaces individuais que compõem uma rede de computadores, na prática, não seguem necessariamente a ordem acima.

No entanto, para utilizar o método de avaliação de disponibilidade, proposto no capítulo anterior, é necessário que os componentes da rede sejam considerados segundo uma ordem crescente de disponibilidades para a geração dos estados de falha mais prováveis da rede (vide capítulo anterior para obter maiores detalhes).

- Não são especificados os tipos de aplicações que executam nas redes;

Todas as configurações topológicas utilizadas nos estudos de caso realizados são apresentadas no apêndice C. Nessas configurações topológicas, os componentes estão numerados segundo a ordem crescente de suas disponibilidades.

A seguir, apresenta-se a análise dos resultados experimentais obtidos nos estudos de casos realizados.

5.2. Estudo do Comportamento dos Limites Superior e Inferior de Disponibilidade em uma Rede de Computadores

Considere a configuração topológica 1 (vide apêndice C). As disponibilidades dos componentes, de tal configuração topológica, estão de acordo com a ordem das disponibilidades da tabela 5.1. Os resultados obtidos, em termos dos limites superior e inferior da nova medida de disponibilidade, ao variar o número de estados de falha mais prováveis (n_{ef}) da rede são listados na tabela, a seguir. Em seguida, apresenta-se o gráfico que relaciona o número de estados de falha mais prováveis aos respectivos limites superior e inferior da nova medida de disponibilidade de redes.

<i>Número de Estados de Falha mais Prováveis</i>	<i>Limites Superior e Inferior da Nova Medida de Disponibilidade</i>
1	1.000000 0.953216
$n_i/5$	0.999789 0.962066
$2n_i/5$	0.999567 0.971381
$3n_i/5$	0.997516 0.978867
$4n_i/5$	0.997294 0.988183
n_i	0.995995 0.994704
$10n_i$	0.995990 0.994913
$20n_i$	0.995971 0.995133
$30n_i$	0.995951 0.995352
$40n_i$	0.995935 0.995574
$50n_i$	0.995925 0.995708
$60n_i$	0.995925 0.995708
$70n_i$	0.995925 0.995708
$80n_i$	0.995925 0.995708
$90n_i$	0.995925 0.995708
$100n_i$	0.995925 0.995708

Tabela 5.2 - Número de Estados de Falha mais Prováveis da rede em Relação aos Limites Superior e Inferior da Nova Medida de Disponibilidade

De acordo com o gráfico da figura 5.1, à medida que o número de estados de falha mais prováveis da rede se aproxima do número total de componentes da rede (n_i), há uma rápida convergência entre os limites superior e inferior da nova medida de disponibilidade. Isto ocorre pois o método de avaliação de disponibilidade utilizado considera que os componentes estão em ordem crescente de disponibilidade, ao gerar os estados de falha mais prováveis da rede. Lembre-se que os componentes estão em ordem crescente de disponibilidade ao serem considerados como entrada para o algoritmo de geração de estados de falha, como discutido no capítulo anterior.

Para os estudos de caso, utiliza-se um valor de $20n_t$ como um valor padrão para o número de estados de falha mais prováveis.

Vale ressaltar que o número de estados de falha mais prováveis necessários para garantir uma determinada diferença entre os limites superior e inferior (uma determinada precisão) depende do número de componentes da rede e de suas respectivas disponibilidades (isto ficará mais claro na seção 5.4, onde a variação das disponibilidades dos componentes é estudada).

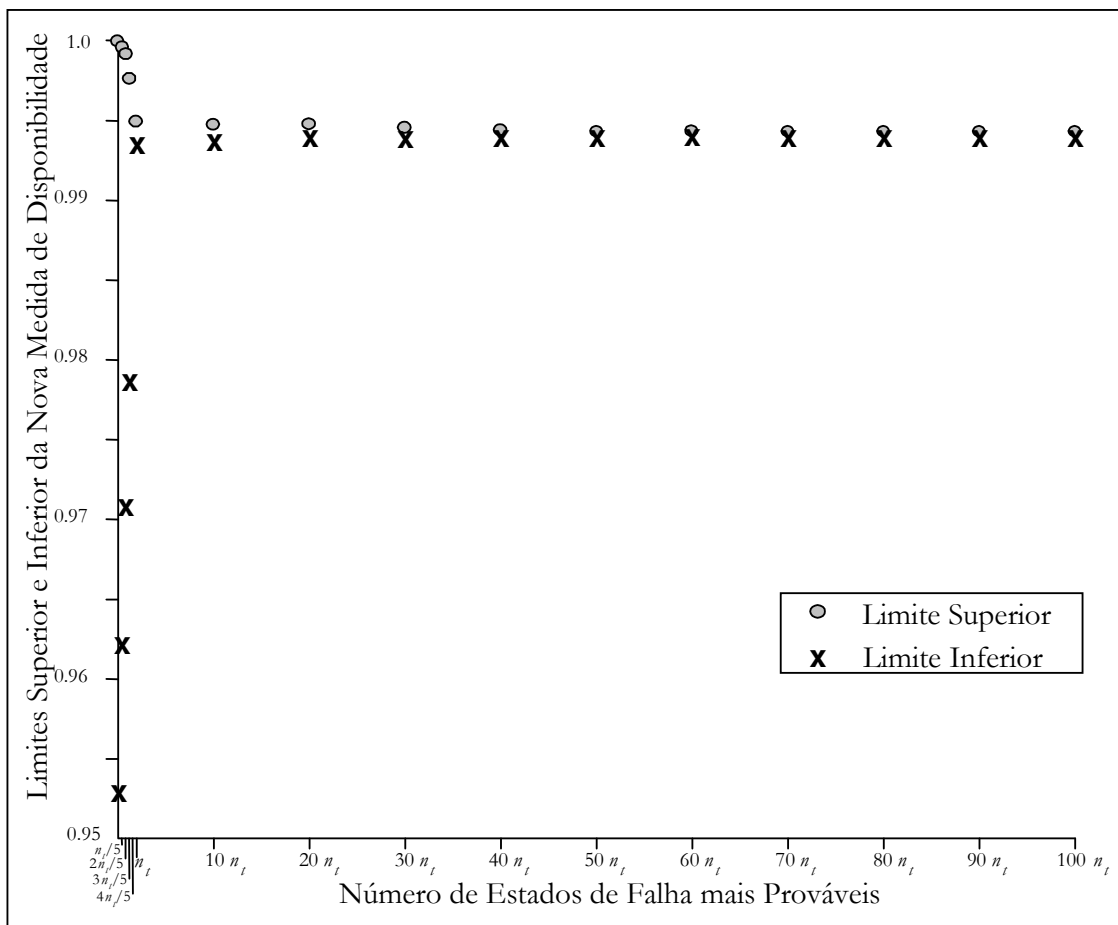


Figura 5.1 - Limites Superior e Inferior da Nova Medida de Disponibilidade x Número de Estados de Falha mais Prováveis

Outra alternativa é implementar o método de avaliação de disponibilidade, proposto no capítulo anterior, com o objetivo de calcular os limites superior e inferior da nova medida de disponibilidade até que a diferença entre eles seja um valor específico (0.00005, por exemplo, é um valor adequado para analisar redes com componentes cujas disponibilidades variam em torno de $0.999x$, uma vez que pode-se ter uma precisão de quatro casas decimais). Entretanto, não se pode definir qual a complexidade

computacional de tal implementação, uma vez que os cálculos dos limites superior e inferior da nova medida de disponibilidade seriam realizados um número de vezes indeterminado para cada configuração topológica particular, de acordo com a precisão requerida e dependendo do número de componentes e de suas respectivas disponibilidades.

Sendo assim, para realizar os estudos de caso, apresentados neste capítulo, considera-se um valor fixo de estados de falha mais prováveis da rede.

5.3. Estudo do Impacto da Adição de Componentes em uma Rede de Computadores

5.3.1. Adição de Clientes

Para estudar o impacto da adição de clientes a uma determinada configuração topológica, no que se refere aos limites superior e inferior da nova medida de disponibilidade, considere as configurações topológicas 2, 3, 4, 5, 6, 7 e 1 (vide apêndice C), nesta ordem. A tabela 5.3 apresenta os limites superior e inferior da nova medida de disponibilidade para cada uma das configurações topológicas analisadas. Em seguida, apresenta-se o gráfico que relaciona o número de clientes aos limites superior e inferior da nova medida de disponibilidade, para tais configurações.

<i>Número de Clientes</i>	<i>Número de Estados de Falha ($m = 20n_i$)</i>	<i>Limites Superior e Inferior da Nova Medida de Disponibilidade</i>
13	800	0.995845 0.995644
18	1000	0.995862 0.995639
23	1200	0.995883 0.995590
28	1400	0.995905 0.995511
33	1600	0.995927 0.995409
38	1800	0.995948 0.995282
43	2000	0.995971 0.995133

Tabela 5.3 - Número de Clientes em relação aos Limites Superior e Inferior da Nova Medida de Disponibilidade

De acordo com os resultados experimentais obtidos, os limites superior e inferior de disponibilidade continuam aproximadamente iguais (observe o gráfico da figura 5.2), ao adicionar clientes. Este é um resultado esperado, uma vez que a adição de clientes não influencia as conectividades parciais consideradas pela nova medida de disponibilidade. O que realmente importa é a disponibilidade de serviços para os clientes adicionais e não a quantidade de clientes adicionados, ao avaliar a disponibilidade das configurações topológicas em questão. Portanto, este resultado assegura que a disponibilidade da infra-estrutura da rede em si está sendo avaliada.

Lembre-se que as conectividades parciais necessárias para que clientes recebam serviços de

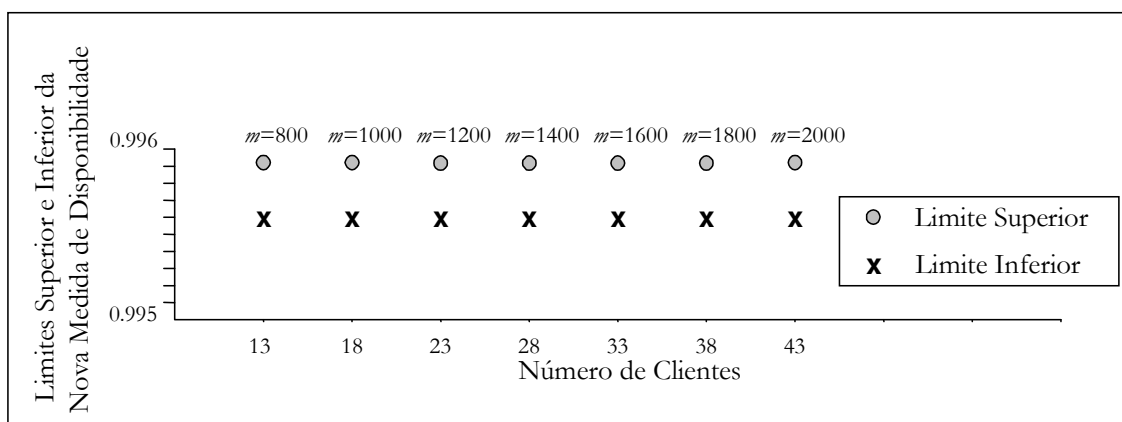


Figura 5.2 - Limites Superior e Inferior da Nova Medida de Disponibilidade x Número de Clientes

uma determinada rede de computadores baseada na arquitetura cliente/servidor em n -camadas são as seguintes:

- Os clientes precisam estar conectados a seus respectivos servidores departamentais;
- Os clientes precisam estar conectados aos servidores corporativos da camada intermediária;
- Os servidores corporativos da camada intermediária precisam estar conectados entre si e os servidores corporativos da camada de dados;
- Os servidores corporativos da camada de dados precisam estar conectados aos servidores corporativos da camada intermediária, no entanto, não precisam estar conectados entre si.

É importante ressaltar que os resultados obtidos em relação à análise da convergência entre os limites superior e inferior da disponibilidade das configurações topológicas utilizadas nos demais estudos de caso seguem a mesma tendência dos resultados expostos acima.

5.3.2. Adição de Servidores Departamentais e Servidores Corporativos

As configurações topológicas 8, 9, 10, 11, 12 e 13 (vide apêndice C) foram utilizadas para estudar o impacto da adição de servidores departamentais e corporativos nos valores dos limites superior e inferior da nova medida de disponibilidade. A tabela 5.4 lista os valores dos limites superior e inferior ao considerar números específicos de servidores departamentais e corporativos. Em seguida, o gráfico para tais valores é apresentado.

<i>Número de Servidores Departamentais e Corporativos</i>	<i>Número de Estados de Falha ($m = 20n_i$)</i>	<i>Limites Superior e Inferior da Nova Medida de Disponibilidade</i>
5	2260	0.994785 0.993720
6	2300	0.994123 0.993031
7	2340	0.993460 0.992340
8	2380	0.992800 0.991653
9	2420	0.992614 0.991438

Tabela 5.4 - Número de Servidores Departamentais e Corporativos em relação aos Limites Superior e Inferior da Nova Medida de Disponibilidade

Observe que o gráfico da figura 5.3.

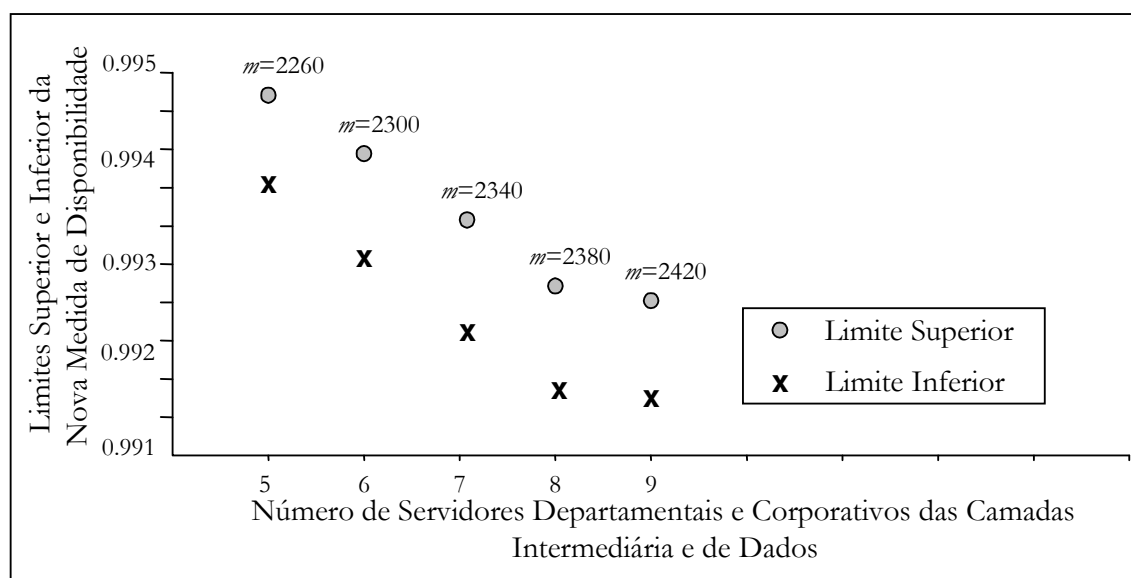


Figura 5.3 - Limites Superior e Inferior da Nova Medida de Disponibilidade x Número de Servidores Departamentais e Corporativos

Ao calcular os limites superior e inferior de disponibilidade, consideram-se as conectividades parciais entre os clientes e os servidores departamentais, como lembrado anteriormente. Sendo assim, é esperado que a adição de servidores departamentais e corporativos cause impacto aos limites superior e inferior da nova medida de disponibilidade, uma vez que as falhas nos servidores departamentais, servidores corporativos e enlaces individuais adicionais serão contabilizados ao calcular os limites superior e inferior da nova medida de disponibilidade, principalmente se não for utilizado algum tipo de redundância física, como é o caso das configurações topológicas analisadas.

Além disso, falhas nos equipamentos de interconexão que conectam os clientes e os servidores departamentais e corporativos também causam impacto aos valores dos limites superior e inferior da nova medida de disponibilidade.

5.4. Estudo do Impacto da Variação da Disponibilidade dos Componentes de uma Rede de Computadores

5.4.1. Variando a Disponibilidade dos Componentes

Ao estudar o impacto da variação das disponibilidades dos componentes aos limites superior e inferior da nova medida de disponibilidade, considere a configuração topológica 14 (vide apêndice C). Para realizar tal estudo, utilizam-se 2880 estados de falha mais prováveis da rede, ou seja, $n_{ef} = 20n_t$. A tabela, a seguir, apresenta os valores dos limites superior e inferior da nova medida de disponibilidade obtidos ao passo que as disponibilidades dos componentes são variadas. Vale ressaltar que se utiliza em cada caso, um valor específico de disponibilidade para todos os componentes da rede.

<i>Disponibilidade</i>	<i>Limites Superior e Inferior da Nova Medida de Disponibilidade</i>
0.9990	0.982344 0.974480
0.9991	0.983891 0.977388
0.9992	0.985483 0.980222
0.9993	0.987123 0.982981
0.9994	0.988811 0.985663
0.9995	0.990547 0.988263

0.9996	0.992334 0.990784
0.9997	0.994172 0.993220
0.9998	0.996062 0.995570
0.9999	0.998003 0.997829

Tabela 5.5 - Limites Superior e Inferior da Nova Medida de Disponibilidade em relação à Variação das Disponibilidades dos Componentes da Rede

De acordo com o gráfico da figura 5.4, à medida que as disponibilidades dos componentes aumentam, os valores dos limites superior e inferior da nova medida de disponibilidade também aumentam. Este resultado é esperado, uma vez que maior a disponibilidade dos componentes, maior é a probabilidade de tais componentes estarem operacionais.

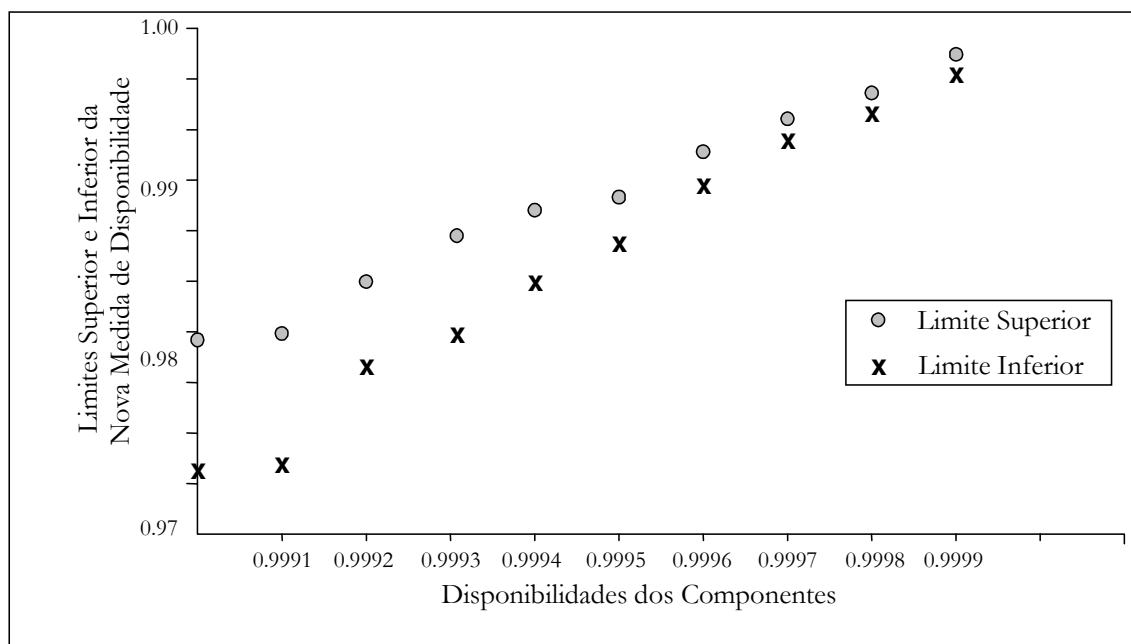


Figura 5.4 - Limites Superior e Inferior da Nova Medida de Disponibilidade x Disponibilidades dos Componentes

A tabela 5.6 apresenta alguns valores de disponibilidade em relação a tempos de quedas por períodos específicos de tempo.

<i>Disponibilidade</i> (%)	<i>Tempo de Queda</i>		
	<i>Mensal</i>	<i>Semanal</i>	<i>Diário</i>
99.9%	43.8 min	10.08 min	1.44 min
99.95%	21.9 min	5.05 min	43.2 s
99.99%	4.4 min	1.0 min	8.7 s

Tabela 5.6 - Disponibilidade em relação aos Tempos de Queda por Períodos Específicos

De acordo com a tabela 5.6, diferenças pequenas entre valores de disponibilidades significam tempos de queda significativos por períodos de tempo específicos.

Em adição, à medida que disponibilidades maiores são consideradas, os limites superior e inferior da nova medida de disponibilidade têm valores mais próximos. Isto advém do fato de que o número de estados de falha considerados é fixo, ou seja, 2880. No entanto, o limite superior inicia em 1.0 para qualquer rede de computadores analisada. Sendo assim, considerar valores mais altos de disponibilidades (isto é, mais próximos de 1.0) para os componentes, significa que os valores dos limites superior e inferior serão valores mais próximos, para um número fixo de estados de falha.

5.4.2. Variando a Disponibilidade dos Enlaces Individuais

Para estudar o impacto da variação das disponibilidades dos enlaces individuais aos limites superior e inferior da nova medida de disponibilidade, considere a configuração topológica 14 (vide apêndice C).

Novamente, utilizam-se 2880 estados de falha mais prováveis da rede, ou seja, $n_{ef} = 20n_t$ para obter os valores dos limites superior e inferior da nova medida de disponibilidade.

Os resultados abaixo, foram obtidos analisando a configuração topológica 14 e considerando que os componentes da rede, exceto os enlaces individuais, têm as disponibilidades (valores fixos) apresentadas na tabela, a seguir.

<i>Componente</i>	<i>Disponibilidade (%)</i>
Clientes	99.95
Equipamentos de Interconexão	99.98
Servidores Departamentais	99.98

Servidores Corporativos da Camada Intermediária	99.98
Servidores Corporativos da Camada de Dados	99.98

Tabela 5.7 - Disponibilidades Consideradas para Estudo de Caso Específico

Os valores dos limites superior e inferior da nova medida de disponibilidade, em relação aos valores das disponibilidades dos componentes são apresentados, a seguir.

<i>Disponibilidade</i>	<i>Limites Superior e Inferior da Nova Medida de Disponibilidade</i>
0.9990	0.991649 0.988716
0.9991	0.992151 0.989478
0.9992	0.992659 0.990240
0.9993	0.993173 0.991002
0.9994	0.993632 0.991344
0.9995	0.994219 0.992527

Tabela 5.8 - Limites Superior e Inferior da Nova Medida de Disponibilidade em relação à Variação das Disponibilidades dos Enlaces Individuais da Rede

A falha de um ou mais enlaces individuais pode acarretar a perda de conectividade parcial ou total da rede. Sendo assim, as disponibilidades dos enlaces individuais causam impacto nos limites superior e inferior da nova medida de disponibilidade. De acordo com o gráfico da figura 5.5, há um aumento dos valores dos limites superior e inferior da nova medida de disponibilidade, à medida que as disponibilidades dos enlaces individuais aumentam.

Uma vez que os enlaces individuais são cruciais para as conectividades parciais consideradas, ao analisar a disponibilidade de redes de computadores é de suma importância que as configurações topológicas possam prover enlaces individuais redundantes e, conseqüentemente, caminhos alternativos. Dada esta importância, apresenta-se um estudo a respeito do impacto da adição de enlaces individuais redundantes em configurações topológicas particulares, na próxima seção.

Observe o gráfico da figura 5.5 que relaciona os valores da tabela 5.6.

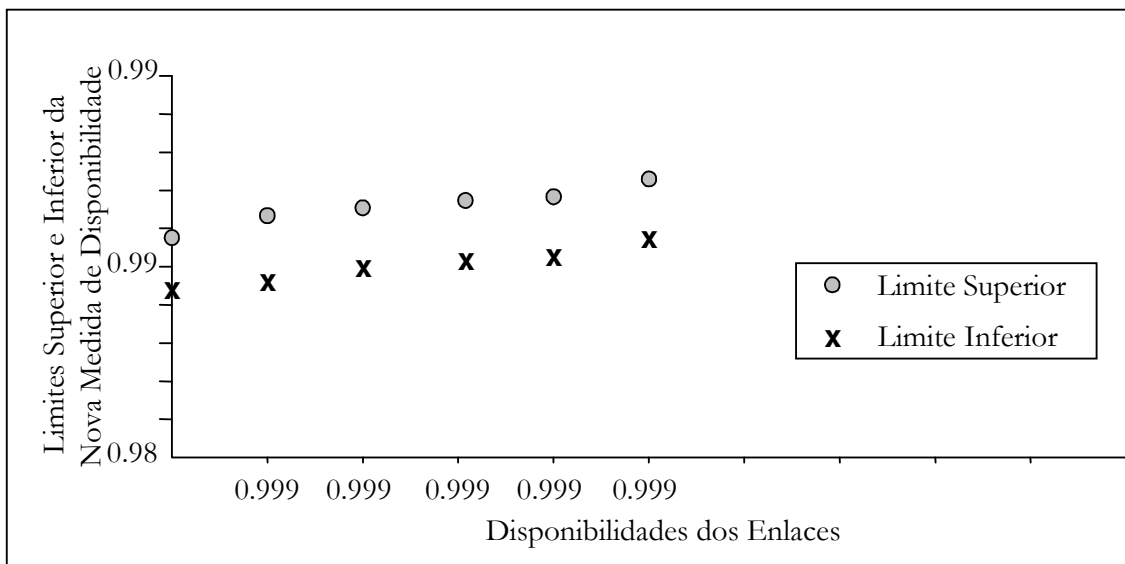


Figura 5.5 - Limites Superior e Inferior da Nova Medida de Disponibilidade x Disponibilidades dos Enlaces

5.5. Estudo do Impacto da Inserção de Redundância de Recursos Físicos em uma Rede de Computadores

Nesta seção, estuda-se especificamente o efeito da inserção de enlaces redundantes a algumas configurações topológicas. Especificamente, utilizam-se as configurações topológicas 15, 16, 17, 18, 19, 20, 21 e 14 (vide apêndice C), nesta ordem. Os limites superior e inferior da nova medida de disponibilidade para tais configurações são apresentados, a seguir.

<i>Número de Enlaces</i>	<i>Número de Estados de Falha ($m = 20n_i$)</i>	<i>Limites Superior e Inferior da Nova Medida de Disponibilidade</i>
65	2580	0.992364 0.991020
66	2600	0.992840 0.991471
67	2620	0.993315 0.991922
66	2600	0.992839 0.991471
67	2620	0.993315 0.991922
69	2660	0.994453 0.993011

73	2760	0.994392 0.992840
78	2880	0.994219 0.992527

Tabela 5.9 - Limites Superior e Inferior da Nova Medida de Disponibilidade em relação ao Número de Enlaces Individuais da Rede

Vale ressaltar que os limites superior e inferior da nova medida de disponibilidade são calculados considerando, além dos enlaces individuais redundantes, que os componentes da rede continuam tendo suas disponibilidades próprias. As disponibilidades dos componentes das configurações topológicas utilizadas estão de acordo com a tabela 5.1. Clientes, servidores departamentais, servidores corporativos e equipamentos de interconexão podem falhar (falhas em hardware ou software - *reboot*, *crash*, etc.) e, portanto, a redundância dos enlaces individuais não vai melhorar a disponibilidade de tais dispositivos.

Observe que para as primeiras 6 configurações topológicas, a adição de enlaces causou aumento nos limites superior e inferior da nova medida de disponibilidade. Especificamente, ao comparar os limites superior e inferior para as configurações topológicas 15 e 20 (vide apêndice C e observe que a configuração 20 é a mesma configuração 15, no entanto, foram adicionados 4 enlaces individuais) pode-se verificar que houve aumentos significativos nos limites superior e inferior da nova medida de disponibilidade. As configurações topológicas 14 e 21 diferem das configurações anteriores, utilizadas neste estudo, uma vez que são adicionados além de enlaces individuais, alguns equipamentos de interconexão. Veja a tabela 5.10.

<i>Disponibilidade</i> (%)	<i>Tempo de Queda</i>		
	<i>Mensal</i>	<i>Semanal</i>	<i>Diário</i>
99.92%	35.04 min	8.06 min	1.15 min
99.94%	26.28 min	6.04 min	51.84 s

Tabela 5.10 - Disponibilidade em relação aos Tempos de Queda por Períodos Específicos

Para algumas aplicações, tais como comércio eletrônico, quedas na rede podem significar perdas de negócios e de lucros para empresas. Sendo assim, redes de computadores que suportam aplicações missão-crítica devem prover caminhos alternativos, a fim de aumentar a disponibilidade

total da rede, e com isso os tempos de queda por períodos específicos podem ser minimizados. Portanto, as diferenças entre os tempos de quedas por períodos específicos para as disponibilidades, supra citadas na tabela, podem ser consideradas significativas.

Na prática, porém, o custo de adicionar redundância pode ser, algumas vezes, proibitivo. Por exemplo, o custo de adicionar um enlace individual em uma rede de computadores compreende o custo do meio físico de transmissão e custo por porta do equipamento de interconexão ao qual ele está conectado.

Por exemplo, atualmente comutadores *Ethernet/Fast Ethernet* típicos têm um custo em torno de \$500 por conexão ao servidor[NPI00], comutadores *Gigabit Ethernet* custam entre \$588 e \$823 por porta [BRUNO-TOLLY00]. Portanto, dependendo das necessidades das aplicações que uma rede de computadores específica suporta, os ganhos obtidos ao adicionar enlaces individuais em relação ao custo requerido ao adicionar redundância podem ser viáveis para a empresa em questão.

Sendo assim, ao projetar redes de computadores com requisitos específicos de disponibilidade é necessário analisar criticamente os ganhos de disponibilidade em relação ao custo de adicionar redundância.

5.6. Estrutura *Collapsed Backbone* versus Estrutura Hierárquica

Collapsed backbone e estrutura hierárquica são duas formas de construir redes de computadores. *Collapsed backbone* é um backbone centralizado, no qual todos os segmentos de rede são interconectados por meio de um equipamento de interconexão [QUINN-ANDRY-HALLER98].

Uma topologia hierárquica¹ compreende módulos (cada um com um objetivo específico diferente). Tais módulos são: camada central (*core*), camada de distribuição e camada de acesso [OPPENHEIMER99]. Adicionalmente, pode-se também considerar uma camada adicional, isto é, a camada de servidores [QUINN-ANDRY-HALLER98].

A camada central concentra os roteadores e comutadores de altos desempenho e disponibilidade e é responsável pela transferência rápida de tráfego entre as outras camadas da hierarquia e, sendo assim, é um camada crítica para a rede como um todo. Portanto, essa camada deve possuir componentes redundantes.

¹ Para obter detalhes a respeito da estrutura hierárquica de uma rede de computadores, consulte [OPPENHEIMER99] e [QUINN-ANDRY-HALLER98].

Na camada de distribuição encontram-se os roteadores e comutadores que implementam políticas, tais como segurança, roteamento, segmentação de tráfego, etc.

A camada de acesso permite que os usuários tenham acesso ao resto da rede, através de equipamentos de interconexão como concentradores e comutadores, tipicamente.

A camada de servidores concentra os servidores corporativos e os equipamentos de comutação e roteamento necessários para acessar o resto da rede.

A seguir, apresenta-se um estudo comparativo entre uma configuração estruturada segundo *collapsed backbone* e segundo uma estrutura hierárquica, a fim de caracterizar o comportamento dos limites superior e inferior da nova medida de disponibilidade ao passo que a mudança entre os tipos de estruturas ocorre. Para tanto, considere as configurações topológicas 22 e 23 (vide apêndice C). Os cálculos dos limites superior e inferior da nova medida de disponibilidade foram realizados considerando os mesmos clientes, servidores departamentais e servidores corporativos.

Número Total de Componentes	Número de Estados de Falha ($m = 20n_i$)	Limites Superior e Inferior da Nova Medida de Disponibilidade
101	2020	0.995013 0.994156
116	2320	0.993129 0.992077

Tabela 5.11 - Limites Superior e Inferior da Nova Medida de Disponibilidade em relação a Mudanças na Estrutura Topológica da Rede

De acordo com a tabela 5.11, a configuração topológica *collapsed backbone* provê maior disponibilidade de serviços aos clientes que a configuração topológica estruturada hierarquicamente.

A conectividade entre os dispositivos da camada de acesso e da camada de servidores depende da conectividade da camada central. No contexto das configurações topológicas estudadas, a mudança de estrutura causou adição de pontos únicos de falha. E uma vez que, ao calcular os limites superior e inferior da nova medida de disponibilidade considera-se os estados de falha mais prováveis da rede, tais pontos únicos de falha causam efeitos em tais limites.

A falha em qualquer um dos equipamentos da camada de distribuição/central pode causar a perda de conectividade entre os clientes e os servidores corporativos. A falha em qualquer um dos equipamentos da camada de acesso ou dos enlaces que conectam a camada de acesso à camada de

distribuição pode afetar a conectividade dos clientes ao resto da rede. A falha no enlace da camada de servidores é crucial para a conectividade entre os servidores corporativos e o resto da rede.

Sendo assim, tais falhas podem afetar as conectividades necessárias para que os clientes estejam recebendo serviços da rede e, conseqüentemente, ao considerar essas falhas, os limites superior e inferior da nova medida de disponibilidade são valores mais baixos para a configuração hierárquica em relação aos valores para a configuração estruturada como *collapsed backbone*.

Apesar dos resultados obtidos neste estudo, deve-se ressaltar que *collapsed backbone* não é escalável e estruturas hierárquicas facilitam mudanças (já que as interconexões são mais simples), minimizam custos (uma vez que a rede é dividida em camadas e cada camada possui um objetivo específico, os dispositivos empregados em cada camada são específicos para realizar uma determinada função) além de facilitar a inserção de redundância às configurações topológicas.

Portanto, projetistas de redes de computadores devem fazer um estudo cuidadoso a fim de escolher a estrutura topológica adequada às necessidades e limitações de uma rede de computadores particular. Precisamente, devem ser consideradas questões referentes a necessidades de redundância e custos *versus* ganhos de disponibilidade.

5.7. Estudo do Comportamento do Tempo de Processamento do Método de Avaliação de Disponibilidade

Nesta seção, estuda-se o comportamento do método de avaliação de disponibilidade, em termos do tempo de processamento necessário para avaliar configurações topológicas particulares.

A seguir, apresentam-se os tempos de processamento em relação ao número total de componentes. Para tal estudo, utilizaram-se as configurações topológicas 2, 6, 1, 8, 13, 18 e 15 (vide apêndice C), nesta ordem.

É importante ressaltar que os valores dos tempos de processamento apresentados no gráfico da figura 5.6 foram obtidos através de estudos de redes de computadores distintas, ou seja, configurações topológicas com diferentes números de clientes, de servidores departamentais e corporativos, de equipamentos e de enlaces individuais. Além disso, utilizou-se um número fixo de estados de falha mais prováveis da rede para a obtenção dos limites superior e inferior da nova medida de disponibilidade, precisamente foram utilizados 2000 estados de falha.

Número Total de Componentes	Tempo de Processamento (segundos)
40	3.35
80	28.90
100	60.24
113	85.28
121	100.74
130	114.17
144	120.56

Tabela 5.12 - Tempo de Processamento em relação ao Tamanho da Rede

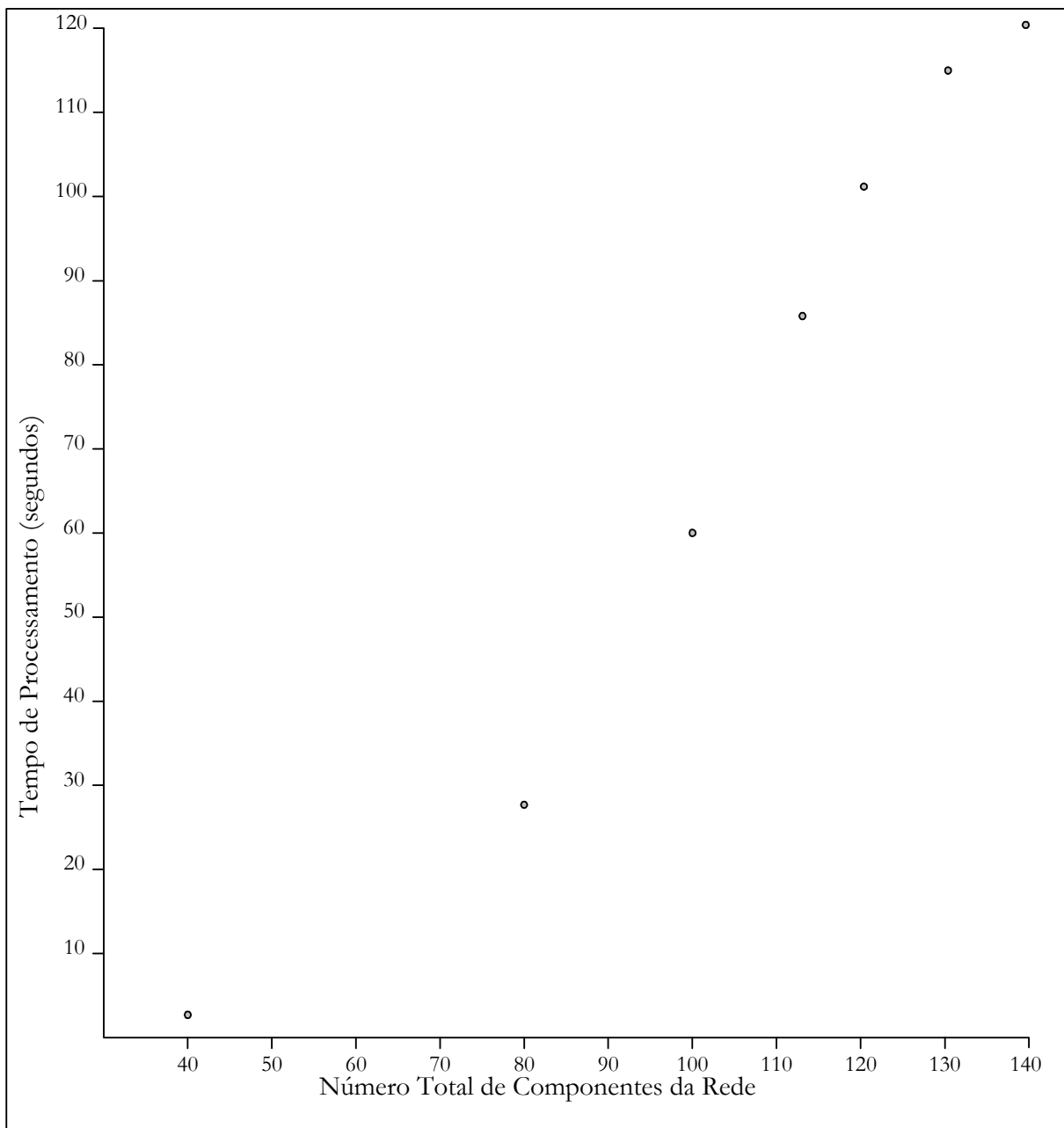


Figura 5.6 - Tempo de Processamento x Número Total de Componentes da Rede

De acordo com o gráfico da figura 5.6, pode-se afirmar que os resultados experimentais obtidos concordam com a análise da complexidade computacional realizada no capítulo anterior. O método de avaliação de disponibilidade tem complexidade computacional polinomial, isto é, $O(n_{ef}n_i^3 + n_{ef}n_i \log n_{ef})$.

Para $n_{ef} = 2000$, tem-se que o número de operações cresce de acordo com $2000n_i^3 + 2000n_i \log 2000$. Pode-se afirmar que $O(2000n_i^3 + 2000n_i \log 2000) = O(n_i^3)$. Sendo assim, os valores da tabela 5.12 estão de acordo com a complexidade computacional do método de avaliação de disponibilidade de redes de computadores baseadas na arquitetura cliente/servidor em n -camadas, proposto no capítulo anterior.

Por exemplo, ao duplicar o valor do número total de componentes (de 40 para 80, por exemplo), espera-se que o tempo de processamento para analisar a rede com 80 componentes seja aproximadamente 8 vezes maior que o tempo necessário para analisar a rede com 40 componentes), o que realmente ocorre, segundo os dados da tabela 5.12 (3.35 e 28.90 segundos, respectivamente).

A seguir, apresentam-se os valores dos tempos de processamento em relação ao número de estados de falha mais prováveis considerados, necessários para calcular os limites superior e inferior da nova medida de disponibilidade de uma configuração topológica particular (configuração topológica 1).

Número de Estados de Falha mais Prováveis	Tempo de Processamento (segundos)
$n_i/5$	0.62
$2n_i/5$	1.20
$3n_i/5$	1.79
$4n_i/5$	2.36
n_i	2.90
$10n_i$	29.48
$20n_i$	58.97
$30n_i$	88.35
$40n_i$	117.03
$50n_i$	146.34
$60n_i$	176.36
$70n_i$	206.22
$80n_i$	236.61
$90n_i$	269.24
$100n_i$	299.92

Tabela 5.13 - Tempo de Processamento em relação ao Número de Estados de Falha mais Prováveis da Rede

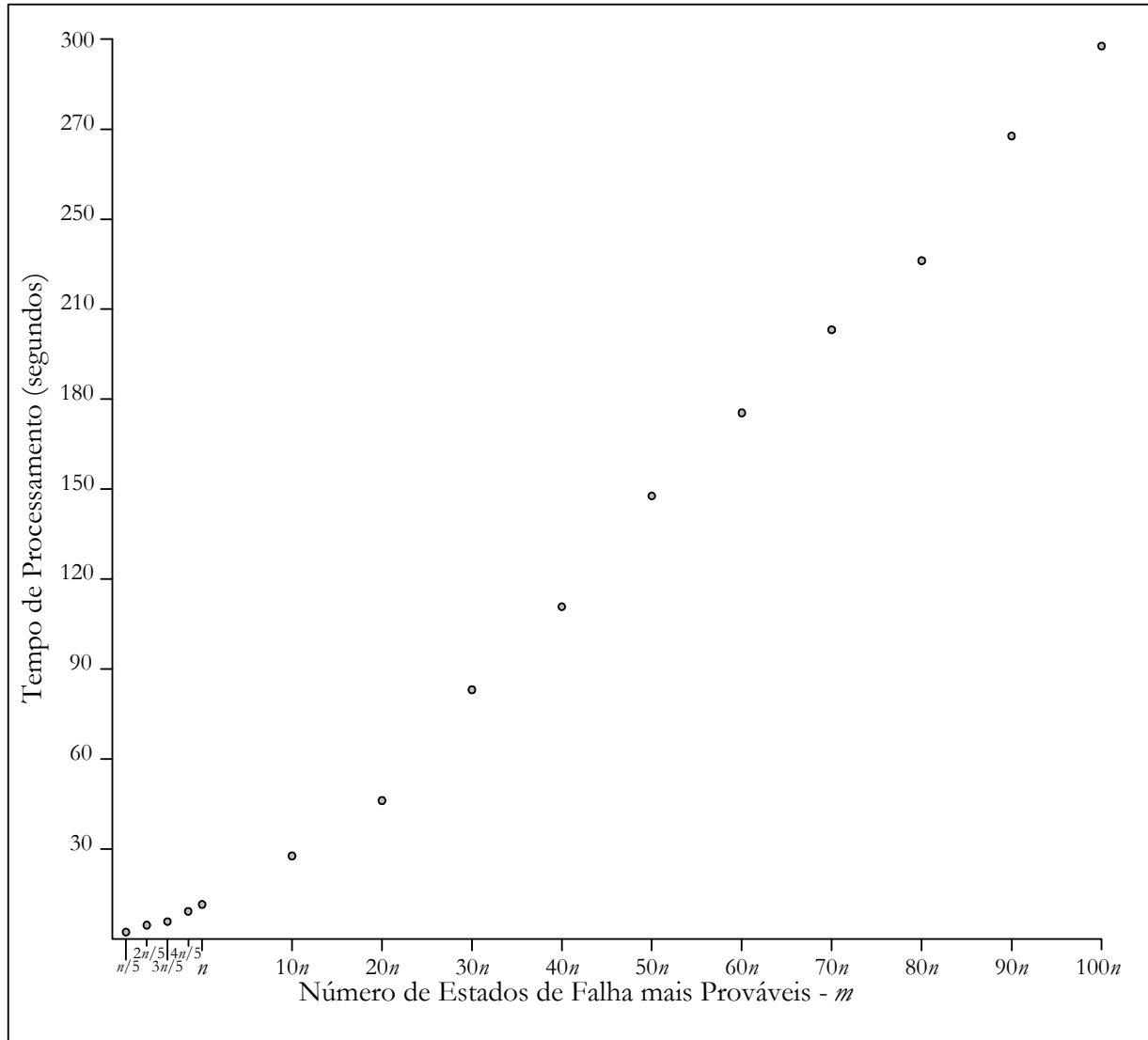


Figura 5.7 -Tempo de Processamento x Número de Estados de Falha mais Prováveis da Rede

Para $n_t=100$, tem-se que o número de operações cresce de acordo com $n_{ef}100 + n_{ef}100\log n_{ef}$. Pode-se considerar que $O(n_{ef}100^3 + n_{ef}100\log n_{ef}) = O(n_{ef}\log n_{ef})$. Por exemplo, ao duplicar o valor do número de estados de falha (de $40n_t$ para $80n_t$, por exemplo), espera-se que o tempo de processamento para analisar a rede, considerando $80n_t$ estados de falha, seja aproximadamente 2.1 vezes maior que o tempo necessário para analisar a rede, considerando $40n_t$ estados de falha. Segundo os dados da tabela 5.13, os tempos de processamento para os números de estados de falha em questão são 117.03 e 236.61 segundos, respectivamente. Observe que a relação entre os tempos de processamento não é exatamente 2.1.

Porém, este fato não significa que a complexidade computacional do método de avaliação de disponibilidade não seja $O(n_{ef}n_t^3 + n_{ef}n_t \log n_{ef})$, uma vez que esta complexidade computacional é assintótica e não se considera, nos estudos dessa seção, valores grandes de n_{ef} .

No próximo capítulo, apresentam-se as conclusões desse trabalho de dissertação, evidenciando ainda, as suas contribuições e propostas para trabalhos futuros relacionados.

CAPÍTULO 6

CONCLUSÃO

Neste capítulo, apresentam-se as conclusões, contribuições e propostas de trabalhos futuros para esta dissertação.

6.1. Conclusões

Os objetivos deste trabalho de dissertação foram atingidos.

No capítulo 3, foi definida uma nova medida de disponibilidade de redes de computadores baseadas na arquitetura cliente/servidor em 3-camadas (ou n -camadas), considerando para tanto, as restrições quanto às conectividades parciais entre clientes e servidores (e entre servidores) impostas para que os clientes recebam serviços completos da rede.

A tabela 6.1 apresenta uma comparação entre os métodos de avaliação de disponibilidade existentes e, em seguida, são apresentadas algumas conclusões provenientes da comparação de tais métodos com o método de avaliação de disponibilidade proposto neste trabalho de dissertação.

Método de Avaliação de Disponibilidade	Medida de Disponibilidade	Técnica de Análise	Modelo Considerado	Complexidade Computacional
[HÄNSLER ET AL.74]	Probabilidade de todos os caminhos entre dois dispositivos estarem operacionais.	Enumeração de estados de falha mutuamente exclusivos.	↻ Todos os dispositivos ou todos os enlaces são considerados suscetíveis a falhas; ↻ Estados	Complexidade computacional exponencial

			operacional e não operacional; ↻ Falhas são consideradas independentes.	
[BALL79]	<ul style="list-style-type: none"> ↻ Probabilidade da rede estar conectada; ↻ Probabilidade de todos os caminhos entre dois dispositivos estarem operacionais; ↻ Probabilidade de todos os dispositivos operacionais estarem se comunicando com um dispositivo específico; ↻ Número de pares de dispositivos que estão se comunicando; ↻ Número de dispositivos que estão se comunicando com um dispositivo específico. 	Particionamento/Decomposição	<ul style="list-style-type: none"> ↻ Dispositivos e enlaces são considerados suscetíveis a falhas; ↻ Estados operacional e não operacional; ↻ Falhas são consideradas independentes. 	Complexidade computacional exponencial.
[KUBAT89]	<ul style="list-style-type: none"> ↻ Probabilidade da rede estar conectada; ↻ Número de pares de dispositivos que estão se comunicando; ↻ Fração de dispositivos se comunicando. 	Simulação/Analítica	<ul style="list-style-type: none"> ↻ Dispositivos e enlaces são considerados suscetíveis a falhas; ↻ Estados operacional e não operacional; ↻ Falhas podem ser dependentes ou independentes. 	Complexidade computacional polinomial.
[AGGARWAL ET AL.75]	Probabilidade da rede estar conectada.	Derivação	<ul style="list-style-type: none"> ↻ Dispositivos e enlaces são considerados suscetíveis a falhas; ↻ Estados operacional e não operacional; ↻ Falhas são 	Complexidade computacional exponencial

			consideradas dependentes.	
--	--	--	---------------------------	--

Tabela 6.1 - Comparação entre Métodos de Avaliação de Disponibilidade de Redes de Computadores

De acordo com a tabela 6.1, alguns métodos de avaliação de disponibilidade foram comparados e pode-se concluir que:

1. As medidas de disponibilidade utilizadas pelos métodos de avaliação de disponibilidade estudados não expressam as restrições de conectividade necessárias para que clientes possam receber os serviços completos de redes de computadores baseadas na arquitetura cliente/servidor em n -camadas;
2. A maioria dos métodos de avaliação de disponibilidade existentes têm complexidade computacional exponencial sendo, portanto, recomendadas para avaliar a disponibilidade de redes de pequeno porte (com um pequeno número de dispositivos). [Kubat89] provê um método de avaliação que tem complexidade computacional polinomial, no entanto, não utiliza uma medida de disponibilidade para o contexto de redes baseadas na arquitetura cliente/servidor em n -camadas.

No capítulo 4, foi formalmente definido um método heurístico de avaliação de disponibilidade que utiliza a nova medida de disponibilidade de redes de computadores baseadas na arquitetura cliente/servidor em n -camadas. Tal método de avaliação, embora não garanta resultados exatos, tem complexidade computacional polinomial, isto é, $O(n_{ef}n_t^3 + n_{ef}n_t \log n_{ef})$, onde n_{ef} é o número de estados de falhas mais prováveis da rede e n_t é o número total de componentes suscetíveis a falhas da rede. Portanto, pode ser utilizado para avaliar a disponibilidade de redes de computadores de maior porte.

A seguir, apresentam-se as principais contribuições deste trabalho de dissertação.

6.2. Contribuições

Redes de computadores corporativas, atualmente, baseiam-se na arquitetura cliente/servidor em 3-camadas (ou n -camadas) e tal fato, como discutido no capítulo 3, impõe algumas restrições quanto à conectividade entre os clientes e servidores (e entre servidores), para que os clientes recebam serviços da rede.

Já que clientes, servidores, equipamentos de interconexão e enlaces individuais podem falhar, é relevante propor uma medida que expresse a fração de clientes que estão recebendo serviços completos da rede, uma vez que falhas podem atingir os clientes apenas parcialmente. No entanto, os métodos de avaliação de disponibilidade existentes, [*BALL-VAN SLYKE*⁷⁷, *HÄNSLER ET AL.*⁷⁴, *KERSHENBAUM-VAN SLYKE*⁷³, *KUBAT*⁸⁹, *MISRA-RAO*⁷⁰, *VAN SLYKE-FRANK*⁷², *VAN SLYKE ET AL.*⁷⁵, *WILKOV*⁷²] entre outros, não consideram tal contexto, como pode ser observado na tabela comparativa 6.1.

Este trabalho propõe um método de avaliação de disponibilidade genérico. Tal método pode ser utilizado para avaliar a disponibilidade de redes de computadores baseadas na arquitetura cliente/servidor em n -camadas, assim como para avaliar a disponibilidade de qualquer tipo de rede de computadores, uma vez que a utilização da matriz de conectividades desejadas possibilita que qualquer tipo de restrições quanto à conectividade entre clientes e servidores (ou entre quaisquer dispositivos) de uma determinada rede de computadores possam ser expressas, ao avaliar a disponibilidade da rede, em questão.

Em adição, o método de avaliação de disponibilidade analisa os estados de falha mais prováveis da rede e estes estados, por sua vez, estão em ordem decrescente de disponibilidade. Sendo assim, tais estados de falha podem dar uma indicação de onde investir na rede a fim de aumentar a sua disponibilidade, uma vez que porções ou pontos críticos podem ser identificados.

Disponibilizar tal método de avaliação de disponibilidade é relevante para analistas de disponibilidade e projetistas de redes de computadores. Hoje em dia, projetistas de redes têm muitas opções de escolha quanto às tecnologias de rede, meios de transmissão e estratégias de interconexão para construir redes de computadores. Sendo assim, os projetistas de rede devem analisar tais opções, considerando os requisitos impostos pelas aplicações e as restrições de custos, a fim de decidir quais estratégias de projeto atendem aos requisitos de uma rede de computadores específica.

O projeto topológico de uma determinada rede de computadores consiste em encontrar uma configuração topológica que atenda às restrições de custo e aos requisitos de disponibilidade impostos por aplicações à rede em questão. O método de avaliação de disponibilidade proposto auxilia nesta tarefa, especificamente, no contexto de redes de computadores baseadas na arquitetura cliente/servidor em n -camadas.

Vale ressaltar que o método de avaliação de disponibilidade proposto não considera questões de projeto, no entanto, pode ser utilizado (e esta é uma das propostas para trabalhos futuros) como parte constituinte de um método de projeto de redes de computadores com requisitos específicos de disponibilidade.

6.3. Trabalhos Futuros

A seguir, listam-se algumas propostas de trabalhos futuros para este trabalho de dissertação.

- Método de projeto de redes de computadores com requisitos específicos de disponibilidade. Tal método de projeto tem por função permitir que projetistas possam construir redes de computadores que garantam os valores específicos de disponibilidade requisitados por suas aplicações próprias, a custos e níveis de desempenho aceitáveis para o projeto em questão. O Método de projeto de redes de computadores deve considerar funções de otimização (por exemplo, minimizar custo e maximizar disponibilidade); mecanismos que aumentam a disponibilidade de redes de computadores, tempos necessários para reconfiguração ou restauração do estado operacional da rede, no evento de falhas; além das metas de projeto (relacionadas a desempenho, por exemplo) para o contexto da rede sendo projetada;
- Tratamento de mecanismos que aumentam a disponibilidade de redes de computadores. Especificamente, aponta-se o tratamento do algoritmo árvore de cobertura, roteamento dinâmico, HSRP (*Hot Standby Routing Protocol*) e DLD (*Deterministic Load Distribution*);
- Particularmente, mecanismos que aumentam a disponibilidade de redes, tais como algoritmo árvore de cobertura, necessitam de um determinado intervalo de tempo para restaurar o estado operacional da rede. Portanto, é interessante que este tempo, bem como o tempo de detecção de falhas, sejam considerados pelo método de avaliação de disponibilidade proposto, ao calcular os limites superior e inferior da nova medida de disponibilidade proposta;

- Este trabalho considerou apenas que os componentes suscetíveis a falhas de uma determinada rede podem estar em um dos estados: estado operacional e estado não operacional. Sendo assim, propõe-se o tratamento do estado intermediário, ou seja, aquele em que um dado componente da rede está operacional, porém opera num nível degradado de desempenho;
- Falhas em componentes da rede afetam seu desempenho. Por tanto, é interessante realizar um estudo do efeito de falhas em componentes sobre a vazão e o atraso na rede;
- Estudo do impacto da utilização de tecnologias de rede diferentes (redes ATM (*Asynchronous Transfer Mode*), redes sem fio, redes *Token Ring*, etc.) sobre os limites superior e inferior da nova medida de disponibilidade de redes proposta.

REFERÊNCIAS BIBLIOGRÁFICAS

- [3COM97A] *SuperStack II Switch 9000 SX*. Product Data Sheet. www.3com.com/products/dsheets/400330a.html, 1997.
- [3COM97B] *OfficeConect Switch 140M*. Product Overview. www.3com.com/solutions/documents/guides/100380.html, 1997.
- [3COM98] *Bulletproofing Networks for High-Availability Computing*. White Paper. www.3com.com/technology/tech-net/white-papers/500693.htm, 1998.
- [3COM00] *Link Builder MSH Multi-Services Hubs*. Product Data Sheet. www.3com.com/products/dsheets/400153.html, 2000.
- [AGGARWAL ET AL.75] K. K. Aggarwal, J. S. Gupta, K. B. Misra. "A Simple Method for Reliability Evaluation of a Communication Systems". IEEE Transactions on Communications, Vol. 23, 1975.
- [AGGARWAL-RAI81] K. K. Aggarwal, S. Rai. "Reliability Evaluation in Computer-Communication Networks". IEEE Transactions on Reliability, Vol. 30, 1981.

- [BALAKRISHNAN ET AL.98] A. Balakrishnan, T. L. Magnanti, P. Mirchandani. "Designing Hierarchical Survivable Networks". Operations Research, Vol.46, 1998.
- [BALL77] M. O. Ball. *Network Reliability Analysis: Algorithms and Complexity*. Tese de Doutorado. Department of Operations Research, Cornell University, 1977.
- [BALL79] M. O. Ball. "Computing Network Reliability". Operations Research, Vol. 27, 1979.
- [BALL80] M. O. Ball. "Complexity of Network Reliability Computations". Networks, Vol. 10, 1980.
- [BALL-VAN SLYKE77] M. Ball, R. M. Van Slyke. "Backtracking Algorithms for Network Reliability Analysis". Annals of Discrete Mathematics, Vol.1, 1977.
- [BARBARA-GARCIA-MOLINA87] D. Barbara, H. Garcia-Molina. "The Reliability of Voting Mechanisms". IEEE Transactions on Computers, Vol. 36, 1987.
- [BAUER ET AL.87] D. Bauer, F. T. Boesch, C. Suffel, R. Van Slyke. "On the Validity of a Reduction of Reliable Network Design to a Graph Extremal Problem". IEEE Transactions on Circuits and Systems, Vol. 34, 1987.
- [BENTLEY86] J. Bentley. *Programming Pearls*. Addison-Wesley Publishing Company, 1986.

- [BIRMAN-JOSEPH87] K. P. Birman, T. Joseph. "Reliable Communication in the Presence of Failures". ACM Transactions on Computer Systems, Vol. 5, 1987.
- [BLAKE79] I. F. Blake. *An Introduction to Applied Probability*. Krieger Publishing Company, 1979.
- [BONAVENTURA80] V. Bonaventura. "Service Availability of Communication Networks". Proceedings 1980 National Telecommunication Conference, 1980.
- [BRUNO-TOLLY00] C. Bruno, K. Tolly. "Competition and Copper Fuel SwitchMetric Gains". Network World, 2000.
- [BUCHANAN-WEAL99] J. Buchanan, E. Weal. *High-Availability Networks*. www.3com.com/nsc/netage/q1_1999/p24.html, 1999.
- [BULTEAU-RUBINO97] S. Bulteau, G. Rubino. *Evaluating Network Vulnerability with the Mincuts Frequency Vector*. Publication Interne No. 1089. Institut de Recherche en Informatique et Systèmes Aléatoires, 1997.
- [CAMPELO-MACULAN94] R. E. Campelo, N. Maculan. *Algoritmos e Heurísticas. Desenvolvimento e Avaliação de Performance*. Editora Universidade Federal Fluminense, 1994.
- [CISCO96] *Using HSRP for Fault-Tolerant IP Routing*. Cisco Documentation. www-fr.cisco.com/univercd/data/doc/cintrnet/ics/icshsrp.htm, 1996.

- [CISCO99A] *HSRP Support*. Product Bulletin #768. www.cisco.com/warp/public/732/100/768-pb.htm, 1999.
- [CISCO99B] *Deterministic Load Distribution*. Product Bulletin #271. www.cisco.com/warp/public/732/102/271-pb.htm, 1999.
- [CISCO99C] *High System Availability*. Product Bulletin. Cisco Systems, 1999.
- [CISCO00] *Prerequisites for High-Availability Infrastructures*. White paper. Cisco Systems, 2000.
- [COCHRANE-HETLEY99] P. Cochrane, D. J. T. Hetley. *Reliability Aspects of Optical Fibre Systems & Networks*. BT Laboratories, 1999.
- [COSTES ET AL.78] A. Costes, C. Landrault, J. C. Laprie. "Reliability and Availability Models for Maintained Systems Featuring Hardware Failures and Design Faults". IEEE Transactions on Computers, Vol. 27, 1978.
- [COTTER-MCGILVREAY99] M. Cotter, J. McGilvrey. *High Availability Industrial Automation Networks*. White Paper. Hirschamann Network Systems, 1999.
- [EDWARDS99] J. Edwards. *3-Tier Server/Client at Work*. John Wiley & Sons, 1999.
- [EVEN79] S. Even. *Graph Algorithms*. Rockville, 1979.
- [FORD ET AL.97] M. Ford, H. K. Lew, S. Spanier, T. Stevenson. *Internetworking Technologies Handbook*. Cisco Press, 1997.

- [FRANK-VAN SLYKE73] H. Frank, R. Van Slyke. "Reliability Considerations in the Growth of Computer Communication Networks". Proceedings of National Telecommunication Conference, 1973.
- [HÄNSLER72] E. Hänsler. "A Fast Recursive Algorithm to Calculate the Reliability of a Communication Network". IEEE Transactions on Communications, Vol. 20, 1972.
- [HAMMOND-O'REILLY86] J. L. Hammond, P. J. P O'Reilly. *Performance Analysis of Local Computer Networks*. Addison-Wesley Publishing Company, 1986.
- [HÄNSLER ET AL.74] E. Hänsler, G. K. McAuliffe, R. S. Wilkov. "Exact Calculation of Computer Network Reliability". Networks, Vol. 4, 1974.
- [HEDDAYA-HELAL96] A. Heddaya, A. Helal. *Reliability, Availability, Dependability and Performability: A User-Centered View*. Technical Report BU-CS-97-11. Boston University, 1996.
- [JALOTE94] P. Jalote. *Fault Tolerance in Distributed Systems*. Prentice Hall, 1994.
- [KERNIGHAN-RITCHIE88] B. W. Kernighan, D. M. Ritchie. *The C Programming Language*. Second Edition. Prentice Hall, 1988.
- [KESHENBAUM-VAN SLYKE73] A. Kershenbaum, R. Van Slyke. "Recursive Analysis of Network Reliability". Networks, Vol. 3, 1973.

- [KE-WANG97] W. Ke, S. Wang. "Reliability Evaluation for Distributed Computing Networks with Imperfect Nodes". IEEE Transactions on Reliability, Vol. 46, 1997.
- [KISHIMOTO97] W. Kishimoto. "Reliable Flow with Failures in a Network". IEEE Transactions on Reliability, Vol.46, 1997.
- [KOPETZ-VERÍSSIMO93] H. Kopetz, P. Veríssimo. *Real Time and Dependability Concepts. Distributed Systems*. ACM Press, 1993.
- [KRÖNSJO87] L. Krönsjo. *Algorithms. Their Complexity and Efficiency*. Second Edition. John Wiley & Sons, 1987.
- [KUBAT86] P. Kubat. "Reliability Analysis for Integrated Networks with Application to Burst Switching". IEEE Transactions on Communications, Vol. 34, 1986.
- [KUBAT89] P. Kubat. "Estimation of Reliability for Communication Computer Networks - Simulation/Analytic Approach". IEEE Transactions on Communications, Vol. 37, 1989.
- [LAM-LI86] Y. F. Lam, V. O. K. Li. "Reliability Modeling and Analysis of Communication Networks with Dependent Failures". IEEE Transactions on Communications, Vol. 34, 1986.
- [LI-SILVESTER83] V. O. K. Li, J. A. Silvester. *An Approach for Studying the Performance of Networks with Unreliable Components*. Technical Report CSI-02-01. University Southern California, 1983.

- [*LI-SILVESTER84*] V. O. K. Li, J. A. Silvester. "Performance Analysis of Networks with Unreliable Components. IEEE Transactions on Communications, Vol. 32, 1984.
- [*MAGALLANES98*] M. D. Magallanes. *PC Reliability Informations*. Modular Industrial Computers, Inc. www.mic.com/pc.htm, 1998.
- [*MCCABE98*] J. D. McCabe. *Practical Computer Networks Analysis and Design*. Morgan Kaufmann Publishers, 1998.
- [*MISRA-RAO70*] K. B. Misra, T. S. M. Rao. "Reliability Analysis of Redundant Networks Using Flow Graphs". IEEE Transactions on Reliability, Vol. 19, 1970.
- [*MYRVOLD96*] W. Myrvold. "Analysis and Synthesis of Reliable Networks". www.csr.wrc.ca/~wendy/rel.html, 1996.
- [*NPI00*] *Fast Ethernet or FDDI? The LAN Administrator's Guide to 100 Mbps Technology*. White Papers. NPI - The Gigabit Ethernet Company, 2000.
- [*OPPENHEIMER99*] P. Oppenheimer. *Top-Down Network Design*. Cisco Press, 1999.
- [*PAPOULIS91*] A. Papoulis. *Probability, Random Variables and Stochastic Processes*. Third Edition. McGraw-Hill, 1991.
- [*PERLMAN92*] R. Perlman. *Interconnections*. Addison-Wesley, 1992.
- [*PIERRE-ELGIBAOU197*] S. Pierre, A. Elgibaoui. "A Tabu-Search Approach for Designing Computer-Network Topologies with Unreliable

Components". IEEE Transactions on Reliability, Vol. 46, 1997.

[PRICE98] R. W. Price. *Ultra-Reliability and the Army After Next*. TRADOC Combat Developments Engineering Division, 1998.

[QUINN-ANDRY-HALLER98] T. Quinn-Andry, K. Haller. *Designing Campus Network*. Cisco Press, 1998.

[ROBERTAZZI90] T. G. Robertazzi. *Computer Networks and Systems. Queuing Theory and Performance Evaluation*. Second Edition. Springer-Verlag, 1990.

[SEYMOUR93] B. Seymour. *MTTR, FailRate, Reliability and Life Testing*. Application Bulletin. Burr-Brown Corporation, 1993.

[SOARES ET AL.95] L. F. G. Soares, G. L. Souza Filho, S. Colcher. *Redes de Computadores. Das LANs, MANs e WANs às Redes ATM*. Editora Campus, 1995.

[SIEWIOREK-SWARZ92] D. P. Siewiorek, R. S. Swarz. *Reliable Computer Systems. Design and Evaluation*. Second Edition. Digital Press, 1992.

[SMITH89] J. D. Smith. *Design and Analysis of Algorithms*. PWS-KENT Publishing Company, 1989.

[SWAMY-THULASIRAMAN81] M. N. S. Swamy, K. Thulasiraman. *Graphs, Networks and Algorithms*. John Wiley and Sons, 1981.

- [TANENBAUM96] A. Tanenbaum. *Networks Computers*. Third Edition. Prentice Hall, 1996.
- [TIZZARD86] K. Tizzard. *C For Professional Programmers*. Ellis Harwood Limited, 1986.
- [TORRIERI94] D. Torrieri. "Calculation of Node-Pair Reliability in Large Networks with Unreliable Nodes". IEEE Transactions on Reliability, Vol. 43, 1994.
- [TRIVEDI86] K. S. Trivedi. *Probability and Statistics with Reliability, Queuing and Computer Science Applications*. Englewood Cliffs - NJ: Prentice Hall, 1982.
- [USHAKOV96] I. A. Ushakov. "Reliability Calculation: To Do or Not To Do?" IEEE Transactions on Reliability, Vol. 45, 1996.
- [VAN SLYKE ET AL. 75] R. Van Slyke, H. Frank, A. Kershenbaum. "Network Reliability Analysis: Part II". Reliability and Fault Tree Analysis (SIAM), 1975.
- [VAN SLYKE-FRANK72] R. Van Slyke, H. Frank. "Network Reliability Analysis: Part I". Networks, Vol. 1, 1972.
- [WALSH96] B. Walsh. *Fault-Tolerant Networking*. Network Computing On line. www.muc.com/netdesign/faultintrob.html, 1996.
- [WILKOV72] R. Wilkov. "Analysis and Design of Reliable Computer Networks". IEEE Transactions on Communication, Vol. 20, 1972.

APÊNDICE A

TEORIA DA CONFIABILIDADE

Neste apêndice, apresentam-se os principais conceitos referentes à confiabilidade de sistemas computacionais. Para tanto, serão discutidas a teoria da confiabilidade e as características referentes à confiabilidade de sistemas computacionais organizados em série, em paralelo e mistos.

Por fim, serão apresentadas medidas tais como o tempo médio entre falhas e o tempo médio de reparo/restauração, de interesse no contexto da teoria da confiabilidade e, finalmente, serão apresentados os conceitos relacionados à disponibilidade de sistemas computacionais.

A.1. Conceitos Básicos

A teoria da confiabilidade está diretamente ligada à avaliação da confiabilidade de sistemas. A maioria dos trabalhos realizados na área de avaliação da confiabilidade de sistemas computacionais baseiam-se na teoria das probabilidades¹, uma vez que a teoria da confiabilidade engloba um contexto onde eventos são aleatórios (falhas podem ocorrer aleatoriamente em sistemas computacionais).

Vale ressaltar que falhas em sistemas computacionais ou em seus componentes estão sujeitos a influências de fenômenos aleatórios ou a mecanismos de mudança. Sendo assim, considerar eventos aleatórios é considerado adequado e, portanto, a teoria das probabilidades pode ser utilizada.

Considerando o contexto da teoria das probabilidades, a maioria dos modelos de confiabilidade de sistemas computacionais baseiam-se em variáveis aleatórias contínuas.

Uma variável aleatória contínua X é uma função que atribui um número real $X(e)$ para cada elemento $e \in E$ (E é o espaço amostral - conjunto de todos os resultados possíveis de um

¹ Os conceitos referentes à teoria das probabilidades estão além do escopo deste apêndice. Para obter informações detalhadas a respeito da teoria das probabilidades, consulte [BLAKE79].

experimento), isto é, atribui um valor real a cada evento do espaço amostral para aquela variável aleatória contínua. Uma variável aleatória contínua requer que o conjunto $\{x \mid X(e) \leq x\}$ seja um evento, para um dado valor x .

Geralmente, em modelos de confiabilidade, três períodos operacionais são de interesse particular: o tempo médio para falhar (MTTF - *Mean Time To Failure*), o tempo médio entre falhas (MTBF - *Mean Time Between Failures*) e o tempo médio para reparo/restauração (MTTR - *Mean Time To Repair*) de um sistema computacional. Tais valores são considerados variáveis aleatórias contínuas, uma vez que não é possível fazer uma previsão de tais valores, antecipadamente.

A.2. Confiabilidade

A confiabilidade de um sistema computacional pode ser considerada como a probabilidade condicional do sistema estar executando sua função de acordo com suas especificações e continuamente durante um determinado intervalo de tempo $[0, t]$, dado que o sistema estava operacional em $t=0$ [BLAKE79].

Esse conceito pode ser generalizado e com isso pode-se afirmar que $C(t, \Delta t)$ é a probabilidade condicional do sistema computacional não falhar no intervalo $[t, t + \Delta t]$, dado que o sistema estava operacional em t , sob determinadas condições ambientais específicas. Valores de confiabilidades são, portanto, probabilísticos e sempre variam entre 0 e 1 e, geralmente, são expressos como frações decimais ou percentagens.

No contexto da teoria da confiabilidade há duas funções de relevância: função de confiabilidade $C(t)$ e a função taxa de falha $c(t)$.

A função taxa de falha é uma função densidade de probabilidade condicional - FDP condicional - tal que $c(t)\Delta t$ é a probabilidade que um sistema computacional falhe durante o intervalo $(t, t + \Delta t)$, dado que o sistema computacional estava operacional em t .

A taxa de falha de um sistema computacional é um valor constante pois, para quaisquer t_1 e t_2 , as taxas $n(t_1)/N(t_1)$ e $n(t_2)/N(t_2)$ dos componentes que falharam durante o intervalo $(t_1, t_1 + \Delta t)$ e $(t_2, t_2 + \Delta t)$, isto é $n(t_1)$ e $n(t_2)$ em relação aos componentes $N(t_1)$ e $N(t_2)$, respectivamente, são constantes e independentes do tempo.

A relação entre $C(t)$ e $c(t)$ é expressa por:

$$C(t+\Delta t) = [1 - c(t) \Delta t]C(t)$$

$$C(t+\Delta t) - c(t) = - C(t) \Delta t c(t)$$

Ou seja,

$$\frac{C(t+\Delta t) - c(t)}{C(t) \Delta t} = -c(t) \quad (\text{Equação 1})$$

Integrando e considerando $C(0) = 1$, obtém-se a seguinte expressão.

$$C(t) = \exp\left(-\int_0^t c(s) ds\right)$$

Considere que a variável aleatória contínua T representa o tempo médio entre falhas de um sistema computacional e assumamos que T tem uma distribuição cumulativa $F_x(t)$. Sendo assim, a confiabilidade do sistema computacional é

$$C(t) = P(X > t) = 1 - F_x(t)$$

Esta equação implica que $C(0) = 1$, ou seja, o sistema computacional está inicialmente operacional e $C(\infty) = 0$, isto é, o sistema computacional não tem tempo médio de vida infinito, isto é, falhará eventualmente.

Considere a equação 1 e já que $C(t+\Delta t) - C(t) = -[F_x(t+\Delta t) - F_x(t)]$, então

$$\frac{C'(t)}{C(t)} = -f_x(t) \quad (\text{Equação 2})$$

Onde $f_x(t)$ é a função densidade de probabilidade de T .

As funções de distribuição de probabilidades e densidade de probabilidades são funções que caracterizam o comportamento de uma variável aleatória contínua.

A distribuição de probabilidades de um componente específico é chamada sua distribuição de falhas. São exemplos de distribuições de falhas as distribuições exponencial negativa, *gamma*, *Weibull* e *log-normal* [BLAKE79].

Freqüentemente, a distribuição de falhas exponencial negativa é usada na avaliação da confiabilidade de sistemas computacionais. Nesse contexto, para uma variável aleatória contínua com distribuição exponencial negativa, a confiabilidade do sistema computacional é:

$$C(t) = e^{-\lambda t}$$

Onde $\lambda = c(t)$ e $f_x(t) = \lambda e^{-\lambda t}$, $t \geq 0$.

Os cálculos realizados anteriormente são referentes a um único sistema computacional isolado. Entretanto, na prática, os sistemas computacionais estão organizados segundo uma combinação de componentes independentes, com características próprias de falhas. Esses componentes podem estar combinados em série, em paralelo ou numa estrutura mista.

Sendo assim, a confiabilidade de um sistema computacional com componentes combinados é o resultado de uma combinação entre as confiabilidades dos componentes. Na

próximas subseções serão definidas formas para calcular confiabilidade de um sistema computacional, consoante a combinação de seus componentes constituintes.

A.2.1. Sistemas Computacionais Organizados em Série

Considere um sistema computacional S composto de n componentes dispostos em série, de acordo com a seguinte figura.

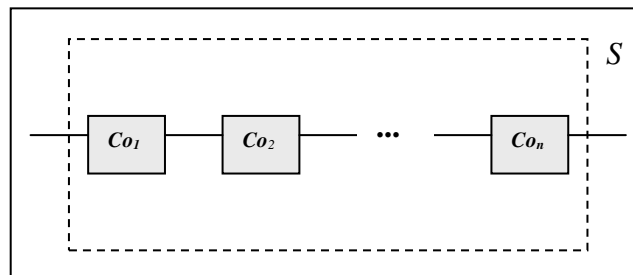


Figura A.1 - Sistema Computacional Organizado em Série

Para o sistema computacional S estar operacional é necessário que os n componentes estejam também operacionais, devido às peculiaridades inerentes a uma organização em série.

Considere que $C_i(t)$ é a probabilidade do componente i estar operacional, num dado tempo t . Sendo assim, a confiabilidade total do sistema computacional S é dada por:

$$C_{série}(t) = C_1(t).C_2(t)...C_n(t)$$

Observe que a confiabilidade total do sistema computacional S é menor que a confiabilidade individual de seus componentes.

Suponha que T_1, T_2, \dots, T_n são tempos em que cada um dos n componentes falham. Portanto, o tempo T para o sistema computacional S falhar é dado por:

$$T_{série} = \min \{T_1, T_2, \dots, T_n\}$$

A.2.2. Sistemas Computacionais Organizados em Paralelo

Considere um sistema computacional S cujos n componentes estão organizados em paralelo, de acordo com a figura A.2.

O sistema computacional S se torna inoperante quando todos os seus componentes falham, devido às características próprias de uma organização em paralelo.

Considere que a probabilidade de um componente S falhar é $1 - C_i(t)$, onde $C_i(t)$ é a sua confiabilidade. Portanto, a confiabilidade do sistema computacional S é:

$$C_{paralelo}(t) = 1 - [1 - C_1(t)].[1 - C_2(t)]...[1 - C_n(t)]$$

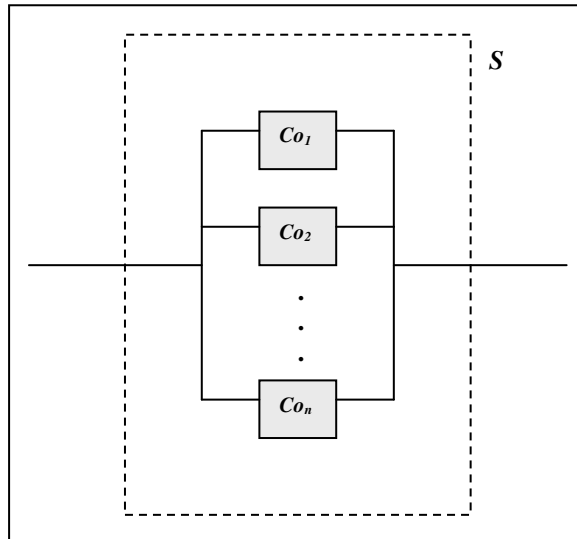


Figura A.2 - Sistema Computacional Organizado em Paralelo

Observe que a confiabilidade do sistema computacional S é maior ou igual à confiabilidade de cada um dos seus componentes.

Sendo assim, supondo que T é o tempo necessário para o sistema computacional S falhar, tem-se:

$$T_{paralelo} = \max \{T_1, T_2, \dots, T_n\}$$

A.2.3. Sistemas Computacionais Organizados em Série e em Paralelo

Na prática, no entanto, os sistemas computacionais estão organizados segundo uma estrutura mista - em série e em paralelo. Um exemplo dessa organização é apresentada na figura A.3.

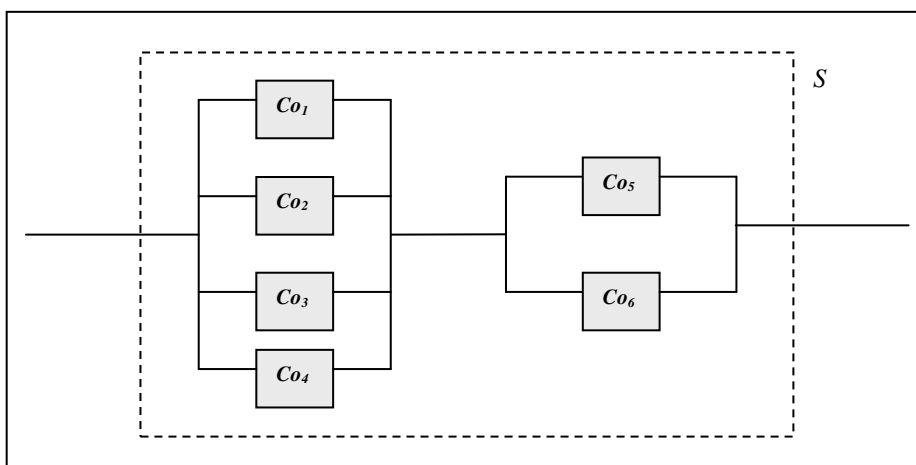


Figura A.3 - Sistema Computacional Organizado em Série e em Paralelo

$C'(t) = 1 - [1 - C_1(t)].[1 - C_2(t)].[1 - C_3(t)].[1 - C_4(t)]$ e $C''(t) = 1 - [1 - C_5(t)].[1 - C_6(t)]$ correspondem às confiabilidades dos sub-sistemas em paralelo. Já que os sub-sistemas estão organizados em série, a confiabilidade do sistema computacional S é:

$$C_{misto}(t) = \{1 - [1 - C_1(t)].[1 - C_2(t)].[1 - C_3(t)].[1 - C_4(t)]\} \cdot \{1 - [1 - C_5(t)].[1 - C_6(t)]\}$$

A seguir, períodos operacionais específicos tais como tempo médio entre falhas e tempo médio de reparo/restauração são discutidos, uma vez que não adianta descrever como um sistema computacional está confiável sem identificar tais períodos, além de especificar as condições ambientais consideradas.

A.3. Tempo Médio para Falhar - MTTF e Tempo Médio entre Falhas - MTBF

Considere, inicialmente, a figura A.4 [SIEWIOREK-SWARZ92].

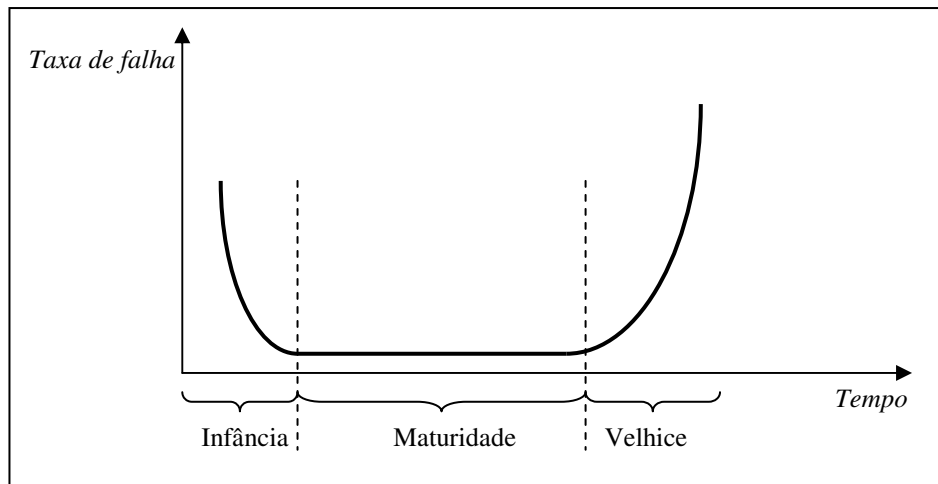


Figura A.4 - Ciclo de Vida de um Sistema Computacional

Observe que o ciclo de vida de um sistema computacional ou componente consiste em três fases: infância, maturidade e velhice. Na primeira fase, a taxa de falha é relativamente baixa e diminui com o tempo decorrido. Na segunda fase a taxa de falha é baixa e constante e na terceira fase a taxa de falha começa a aumentar novamente e rapidamente com o tempo.

MTTF é o inverso da taxa de falha na fase de maturidade do ciclo de vida de um componente ou sistema computacional. Seja λ (considerando a distribuição exponencial negativa) a taxa de falha de um dado sistema computacional ou componente, então

$$MTTF = \frac{1}{\lambda}$$

Confiabilidade está diretamente relacionada ao MTTF. Quando o MTTF aumenta a confiabilidade também aumenta, e vice-versa.

MTBF é o tempo médio que um componente ou sistema computacional executa sem falhas, após ter sido configurado, reparado ou restaurado. Sendo assim, este tempo corresponde à média de tempo transcorrido entre uma falha e outra, ou seja, $MTBF = MTTF + MTTR$.

A.4. Tempo Médio para Reparo - MTTR

Na prática, quando um ou mais componentes de um sistema computacional falham, estes são trocados ou reparados. Sendo assim, pode-se identificar dois estados para um sistema computacional: operacional ou em falha (sob reparo ou ainda não foi restaurado). A transição entre os estados ocorre segundo a figura abaixo.

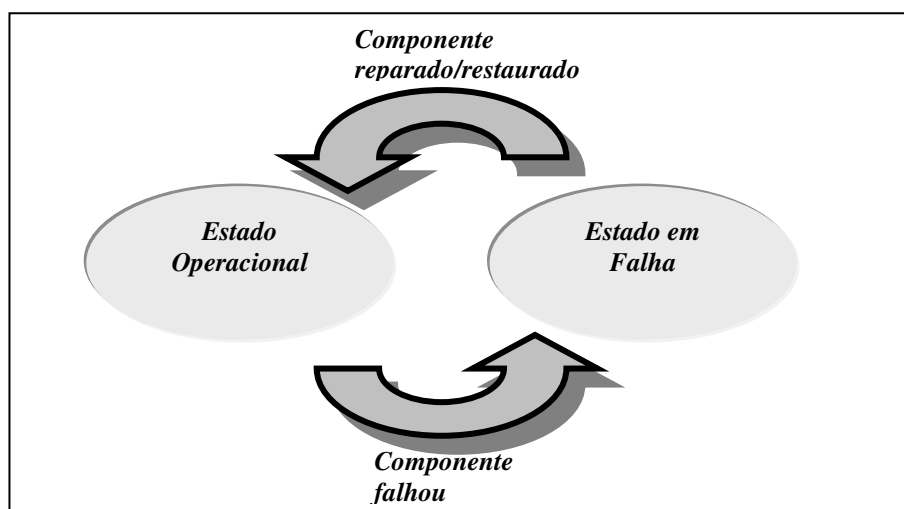


Figura A.5 - Estados de um Sistema Computacional

MTTR é o tempo médio necessário para reparar ou restaurar um sistema computacional ou componente em falha. Sendo assim, o MTTR engloba desde o tempo em que a falha ocorre até o tempo em que o componente ou sistema computacional é reparado/restaurado e, conseqüentemente, retorne ao estado operacional.

Além disso, MTTR é o inverso da taxa de reparo/restauração. Seja μ a taxa de reparo/restauração de um dado sistema computacional ou componente, então

$$MTTR = \frac{1}{\mu}$$

A.5. Disponibilidade

Para sistemas computacionais que podem ser reparados, a função de disponibilidade $D(t)$ é adequada para descrever o comportamento de um sistema computacional. A disponibilidade

$D(t)$ é a probabilidade do sistema estar operacional (de acordo com suas condições funcionais desejadas), no instante de tempo t [BLAKE79].

A disponibilidade $D(t)$ difere de $C(t)$, uma vez que as falhas que ocorreram antes do tempo t não causam impacto na disponibilidade do sistema computacional. Vale ressaltar que em sistemas computacionais sem reparo é adequado utilizar a confiabilidade e esta, por sua vez, é a probabilidade do tempo de vida do sistema computacional exceder t .

Para uma taxa de falha λ e uma taxa de reparo/restauração μ constantes, a disponibilidade no estado permanente de um sistema computacional ou componente pode ser expresso como:

$$\text{Disponibilidade} = \frac{\mu}{\lambda + \mu}$$

APÊNDICE B

ARQUITETURA CLIENTE/SERVIDOR EM N-CAMADAS

Neste apêndice, apresenta-se uma visão geral a respeito da arquitetura cliente/servidor em n -camadas [EDWARDS99]. Inicialmente, são evidenciadas as diferenças entre a arquitetura cliente/servidor original - em duas camadas e a arquitetura em n -camadas. Em seguida, descrevem-se as vantagens de empregar a arquitetura cliente/servidor em n -camadas.

B.1. Arquitetura Cliente/Servidor em 2 Camadas

A arquitetura cliente/servidor em 2 camadas consiste nas seguintes camadas.

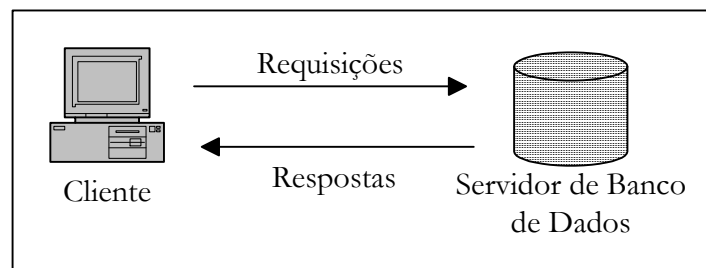


Figura B.1- Arquitetura Cliente/Servidor em 2 Camadas

Vale salientar que todo o processamento é realizado no cliente e este, por sua vez, acessa diretamente os servidores de bancos de dados que funcionam como repositórios de dados. Tal arquitetura possui algumas desvantagens. A seguir, enumeram-se algumas delas:

1. É necessária uma conexão de cada cliente para o servidor de banco de dados. Esta solução não é escalável, uma vez que, cada conexão particular requer tempo de processamento e memória. Em adição, à medida que o número de conexões aumenta, o desempenho do servidor de banco de dados sofre degradação;
2. É necessário instalar um software cliente específico, implementado sob medida, em cada cliente, a fim de que este possa ter acesso ao servidor de banco de dados. Esse fato gera

problemas de gerenciamento. Por exemplo, suponha que há uma nova versão de uma determinada aplicação e é necessário instalá-la em centenas ou até milhares de clientes. Tal instalação não é instantânea e, sendo assim, haverá momentos nos quais as duas versões estarão executando.

Problemas como os anteriormente descritos, aliados ao contexto atual da maioria das empresas, onde há evidência do aumento da demanda por transferência de informações, transações comerciais e necessidades de desenvolver aplicações escaláveis e distribuídas torna necessário empregar um novo tipo de arquitetura cliente/servidor: a arquitetura cliente/servidor em n -camadas.

B.2. Arquitetura Cliente/Servidor em n -Camadas

Nesse novo contexto, o lado servidor foi dividido em duas camadas: camada intermediária e camada de dados. Para uma arquitetura em três camadas, tem-se a estrutura apresentada na figura B.2.

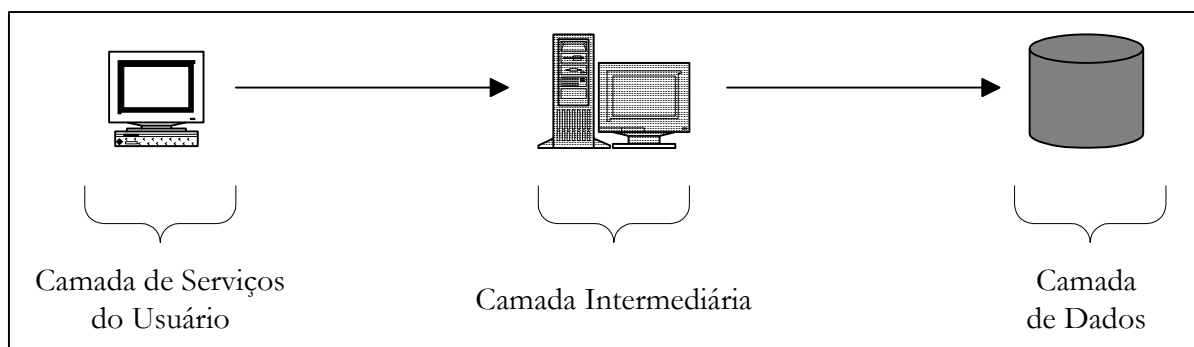


Figura B.2 - Arquitetura Cliente/Servidor em n -Camadas

Na arquitetura cliente/servidor em n -camadas, a comunicação entre os componentes das camadas segue estas restrições: cada cliente deve acessar os seus servidores departamentais próprios e os servidores corporativos da camada intermediária. Os servidores corporativos da camada intermediária devem ter acesso uns aos outros e aos servidores corporativos da camada de dados. Estes últimos, por sua vez, devem acessar apenas a camada intermediária e não têm acesso uns aos outros, tipicamente.

Observe que o cliente não tem acesso direto à camada de dados e que a maioria do processamento passa a ser realizado na camada intermediária, onde estão armazenadas as aplicações. A camada intermediária pode ser implementada em várias camadas, por isso a denominação arquitetura cliente/servidor em n -camadas. Uma das razões para utilizar a camada intermediária é compartilhar recursos caros e escassos utilizados pelos clientes, a exemplo da

economia computacional e de memória ganha com a eliminação de conexões diretas entre clientes e servidor de banco de dados¹.

Uma vez que a maioria do processamento está sendo realizado na camada intermediária², nos clientes estão os folheadores, isto é, software para visualização, necessários para acessar as aplicações contidas na camada intermediária. Quando ocorrem mudanças (por exemplo, novas versões de uma determinada aplicação), a atualização deve ser realizada apenas na camada intermediária sem a necessidade, portanto, de serem repassadas para todos os clientes.

Adicionalmente, os clientes não precisam ter grande poder computacional, uma vez que não é necessária a instalação de software cliente específico. Sendo assim, uma alternativa são os chamados *thin clients*. *Thin client* é um dispositivo com características físicas e de configuração simples, que garantem facilidade de manutenção e permitem o acesso às aplicações em servidores, por parte das aplicações. Em tais dispositivos nenhum ou o mínimo de instalação de software é requerida.

¹ Há também outros motivos, tais como balanceamento de carga e *failover* automático, entre outros motivos.

² É possível ter algum processamento na camada de dados, através de *stored procedures*.

APÊNDICE C

CONFIGURAÇÕES TOPOLÓGICAS UTILIZADAS NOS ESTUDOS DE CASO

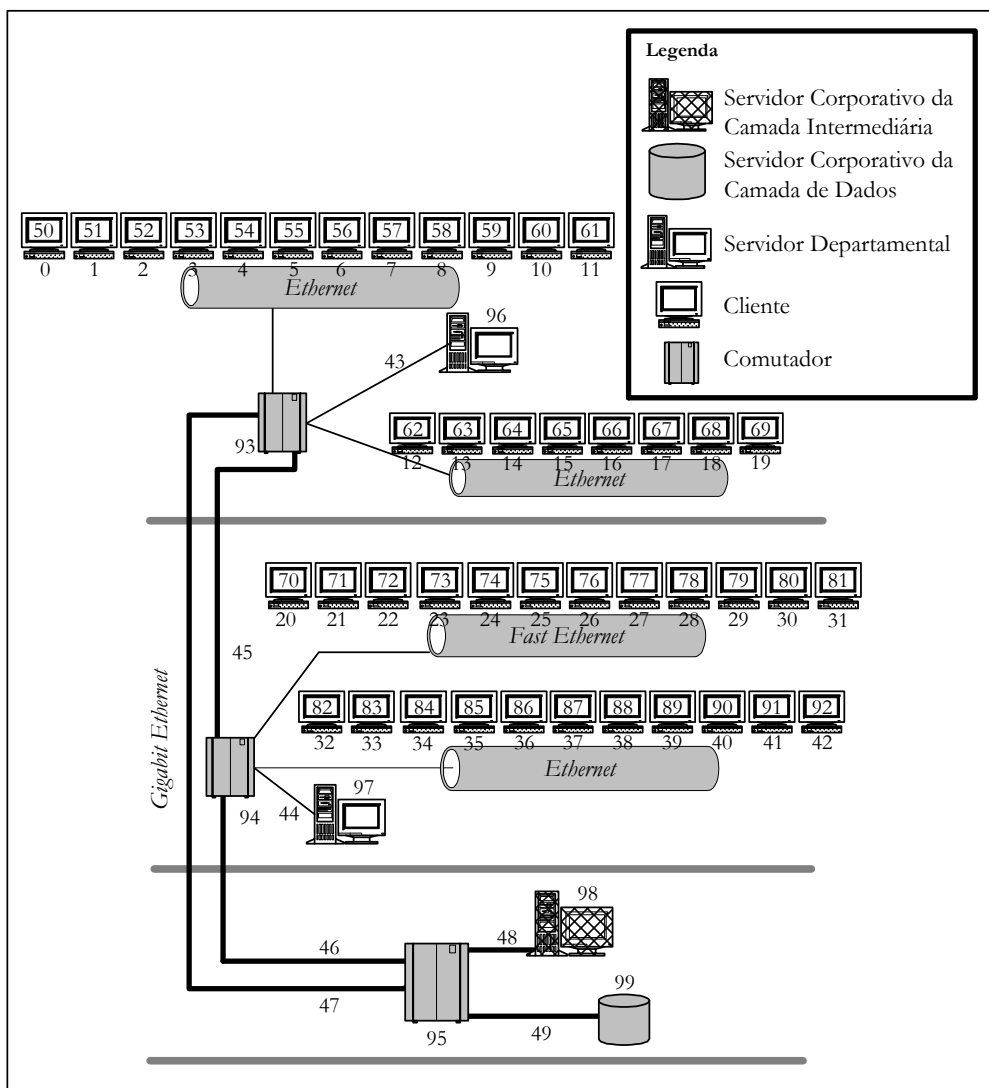


Figura C.1 - Configuração Topológica 1

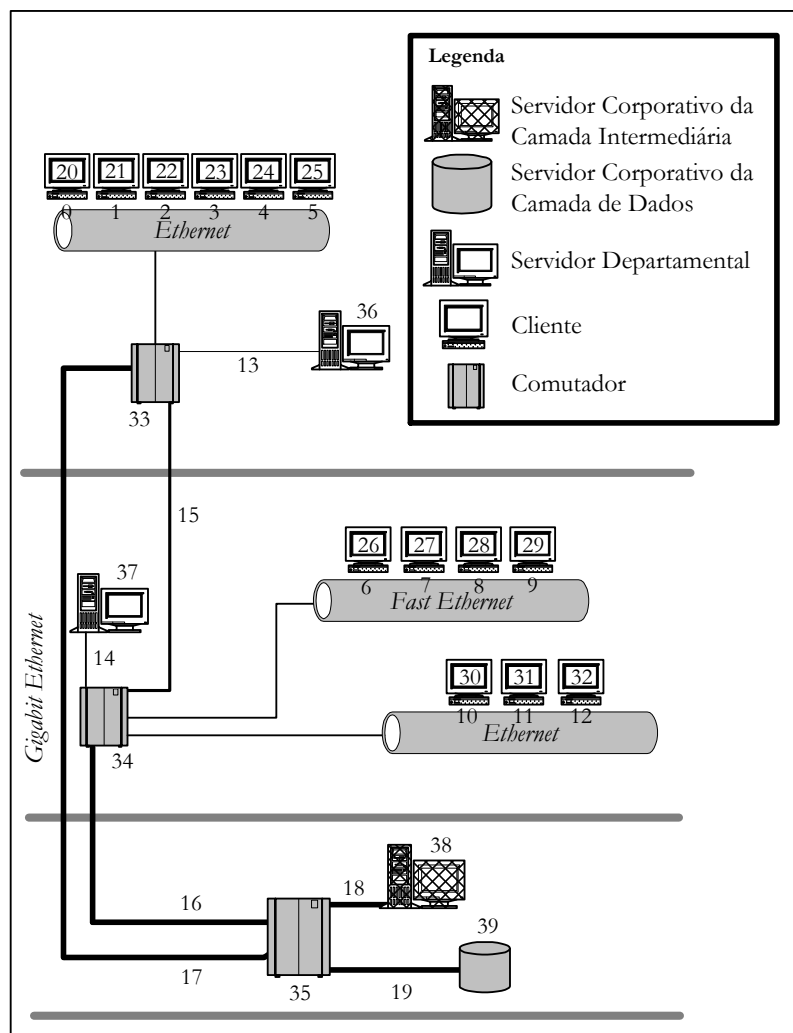


Figura C.2 - Configuração Topológica 2

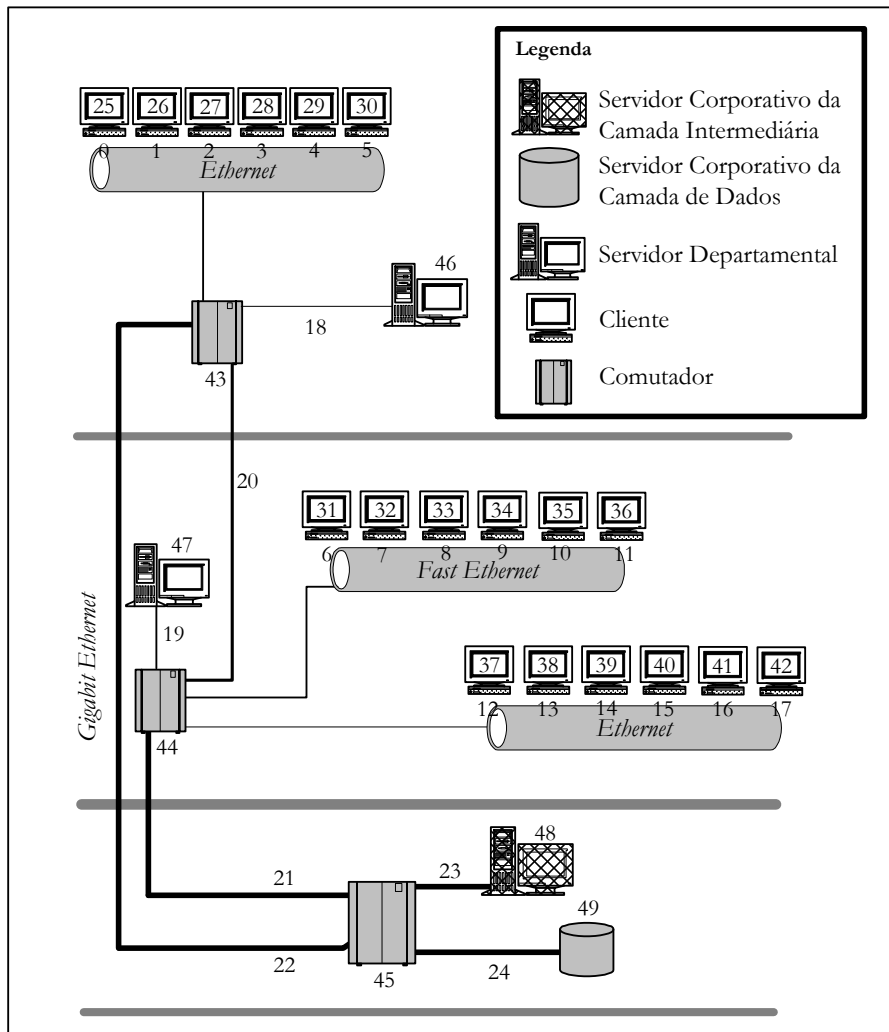
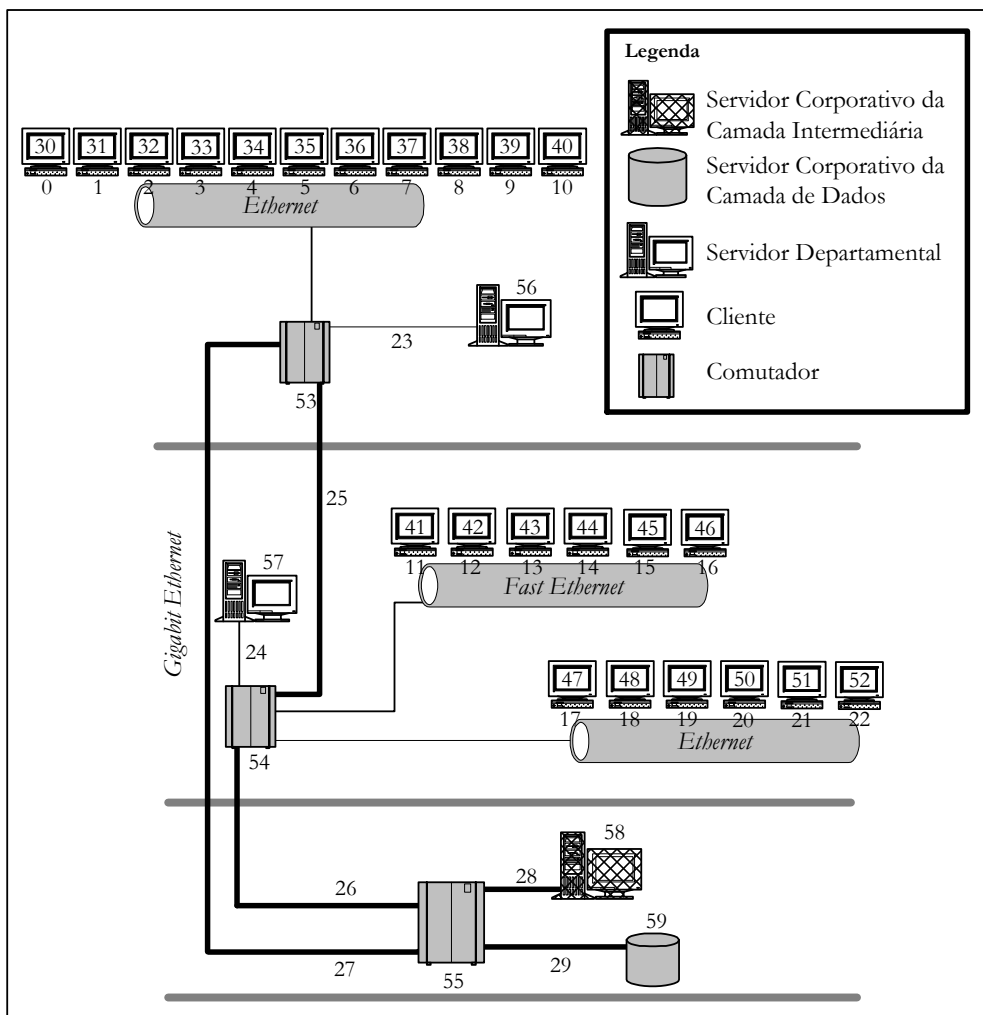


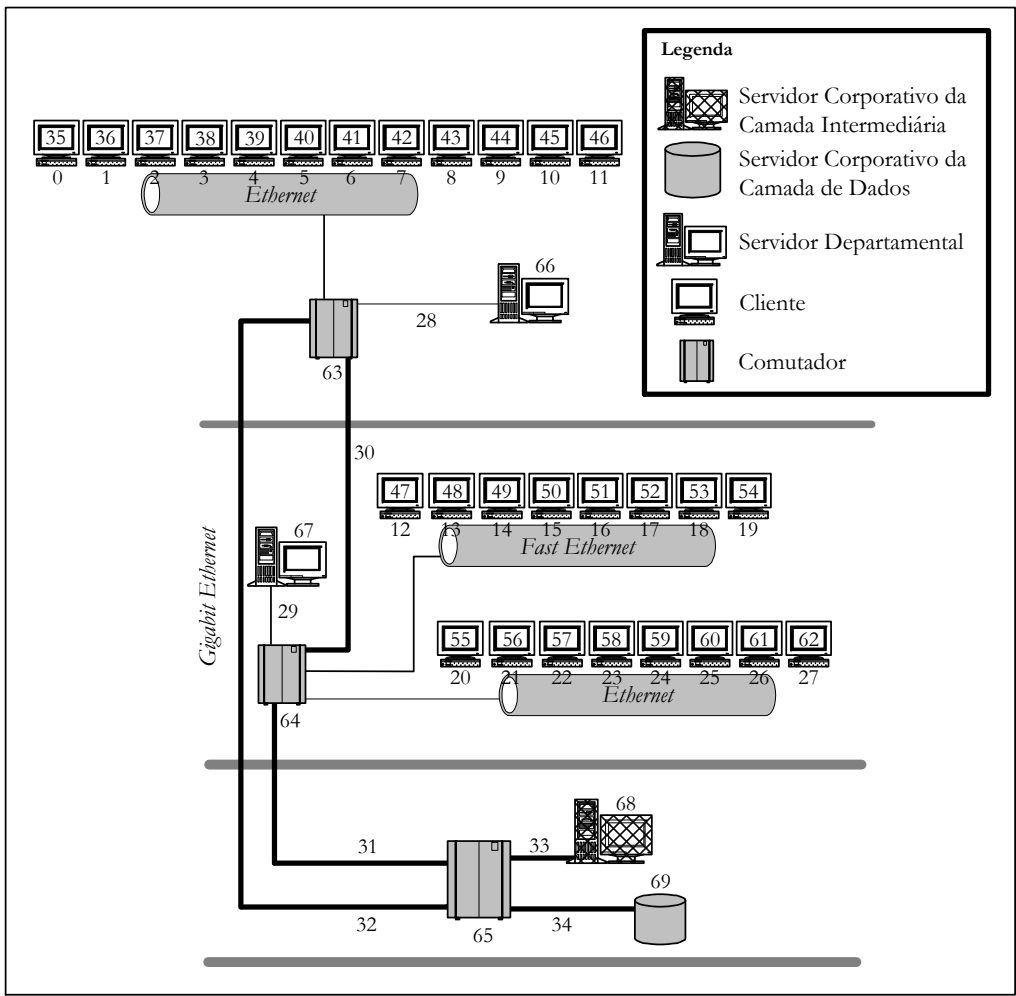
Figura C.3 - Configuração Topológica 3



Legenda

	Servidor Corporativo da Camada Intermediária
	Servidor Corporativo da Camada de Dados
	Servidor Departamental
	Cliente
	Comutador

Figura C.4 - Configuração Topológica 4



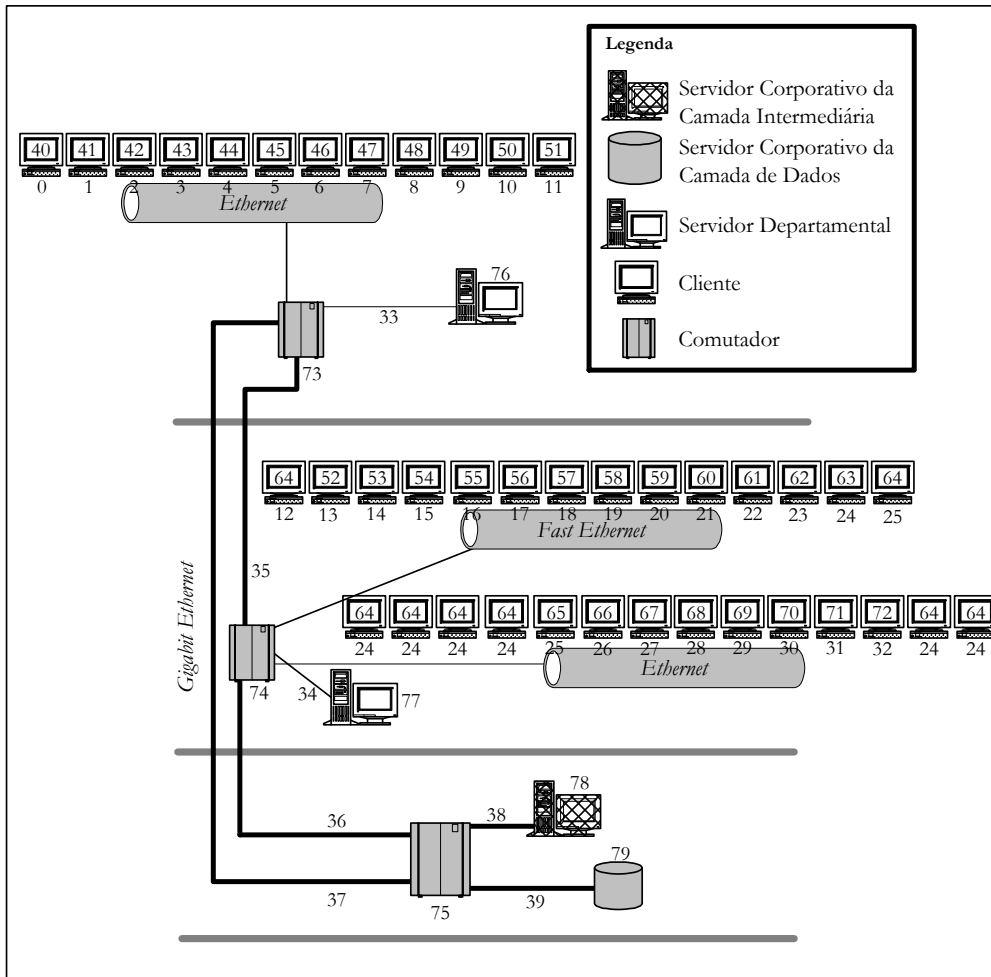


Figura C.6 - Configuração Topológica 6

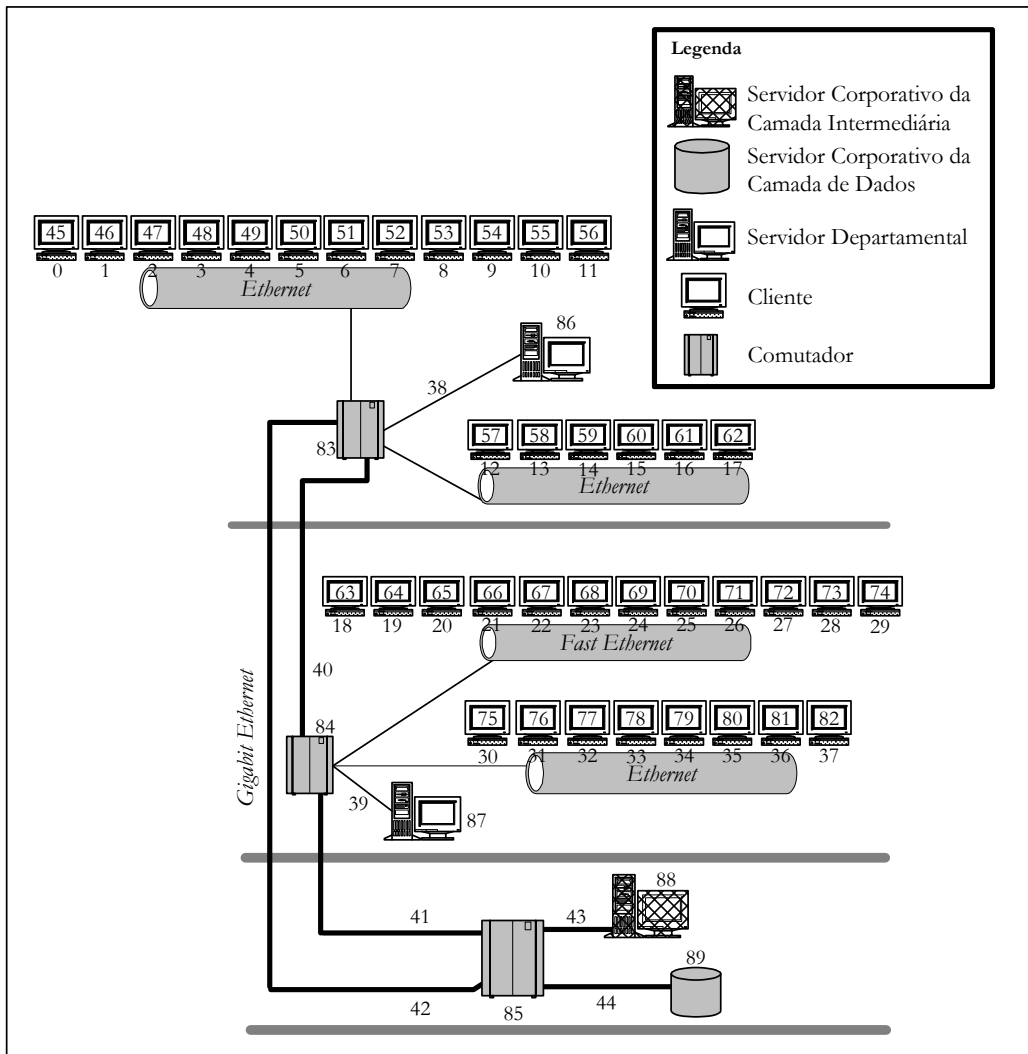


Figura C.7 - Configuração Topológica 7

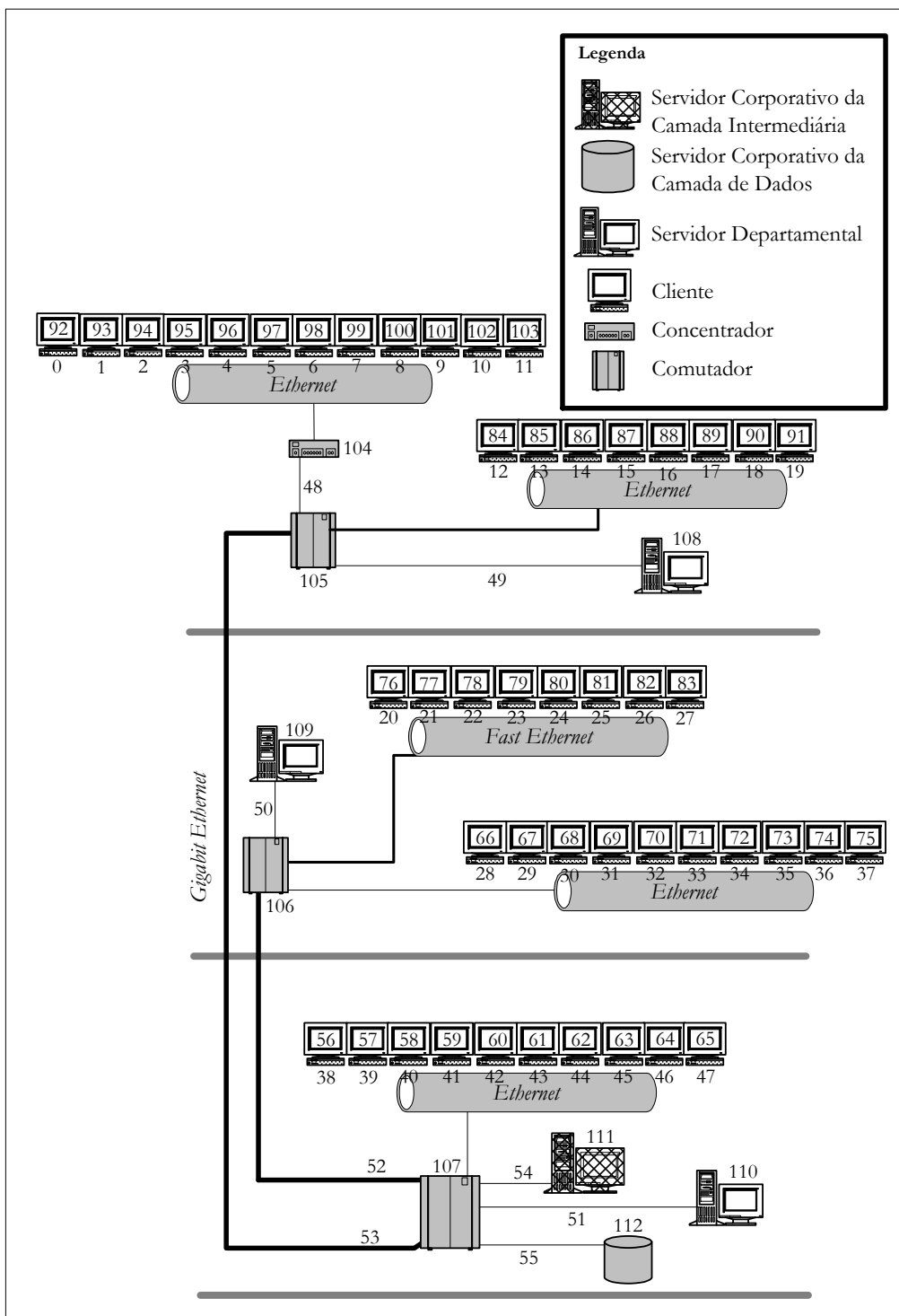


Figura C.8 - Configuração Topológica 8

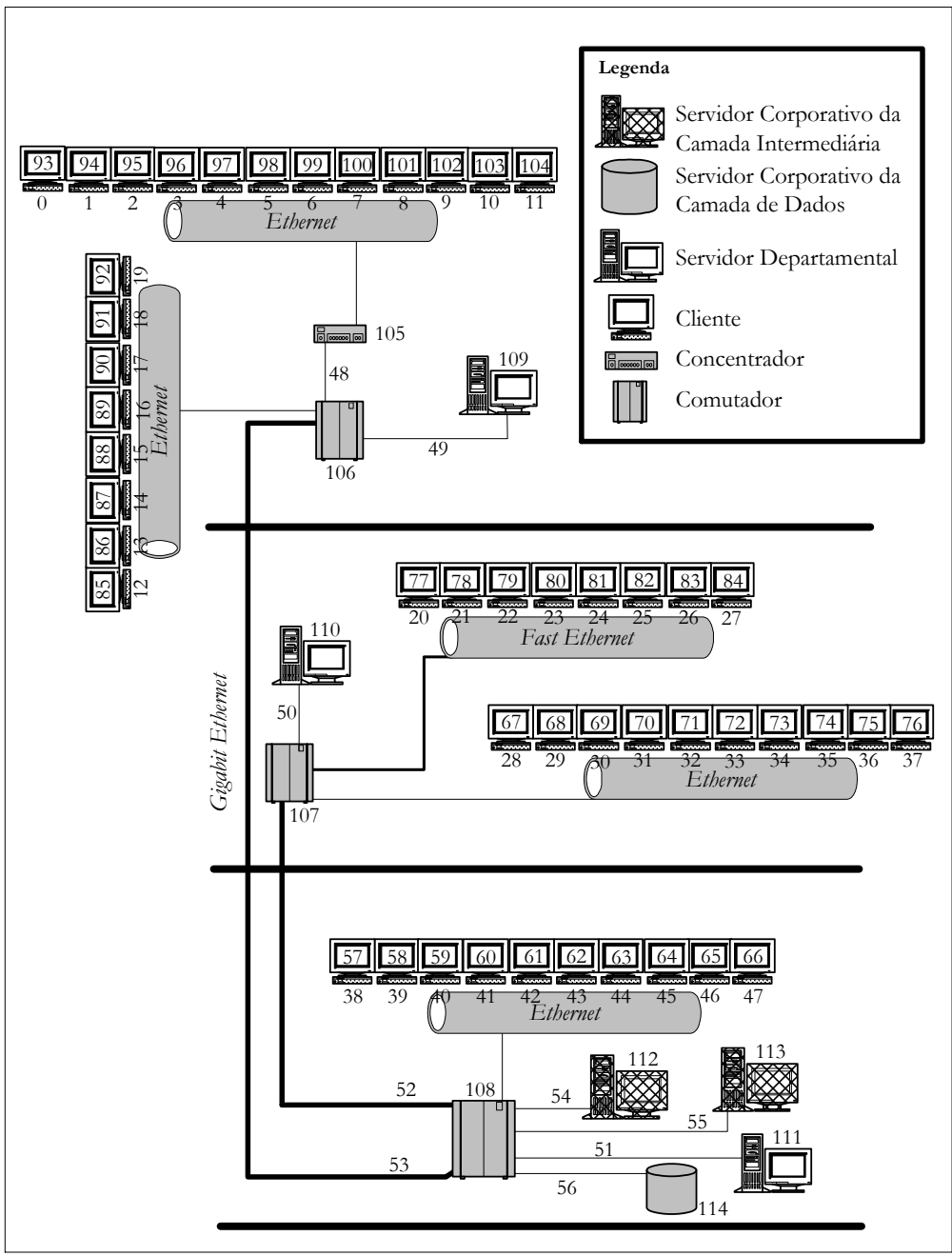


Figura C.9 - Configuração Topológica 9

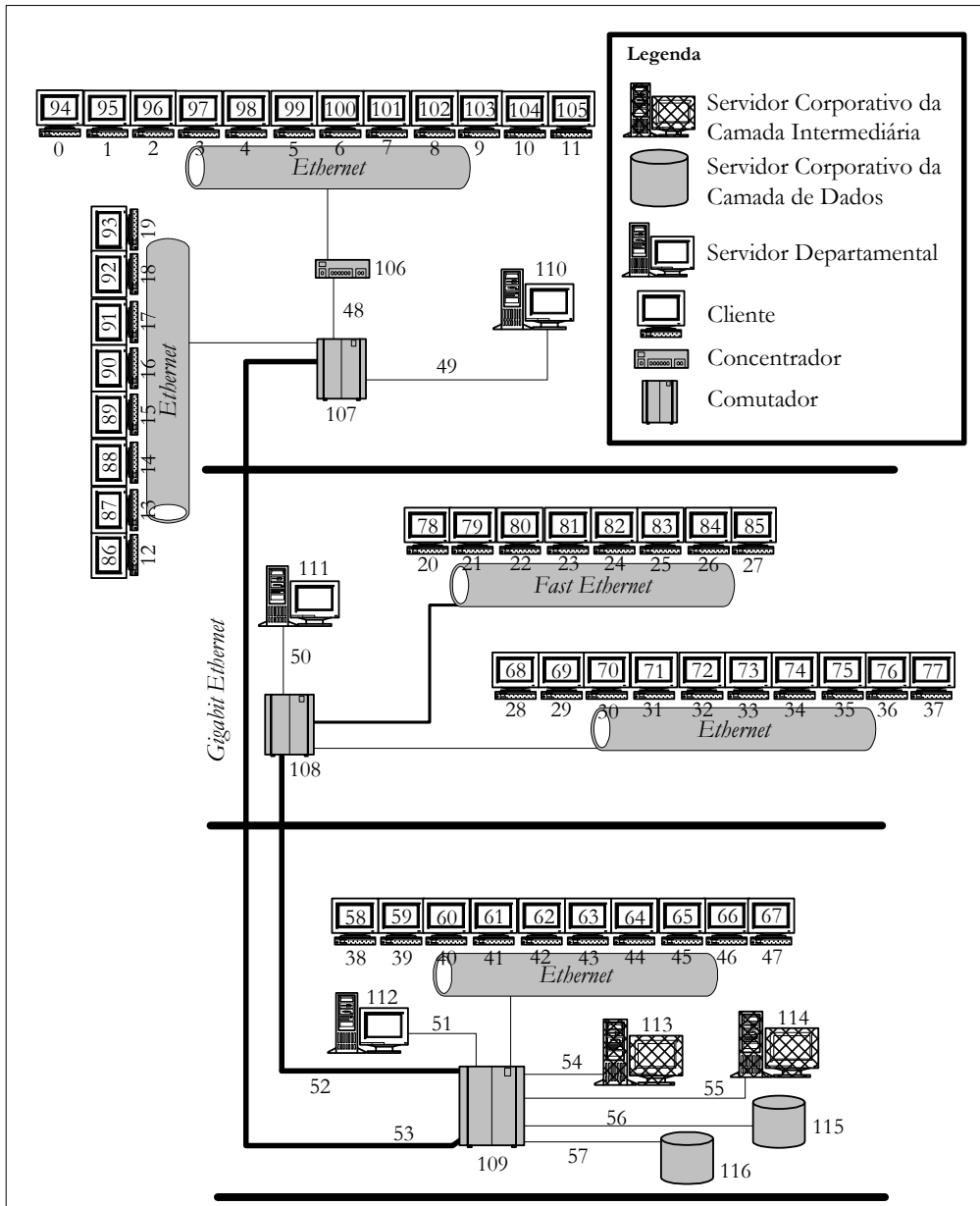


Figura C.10 - Configuração Topológica 10

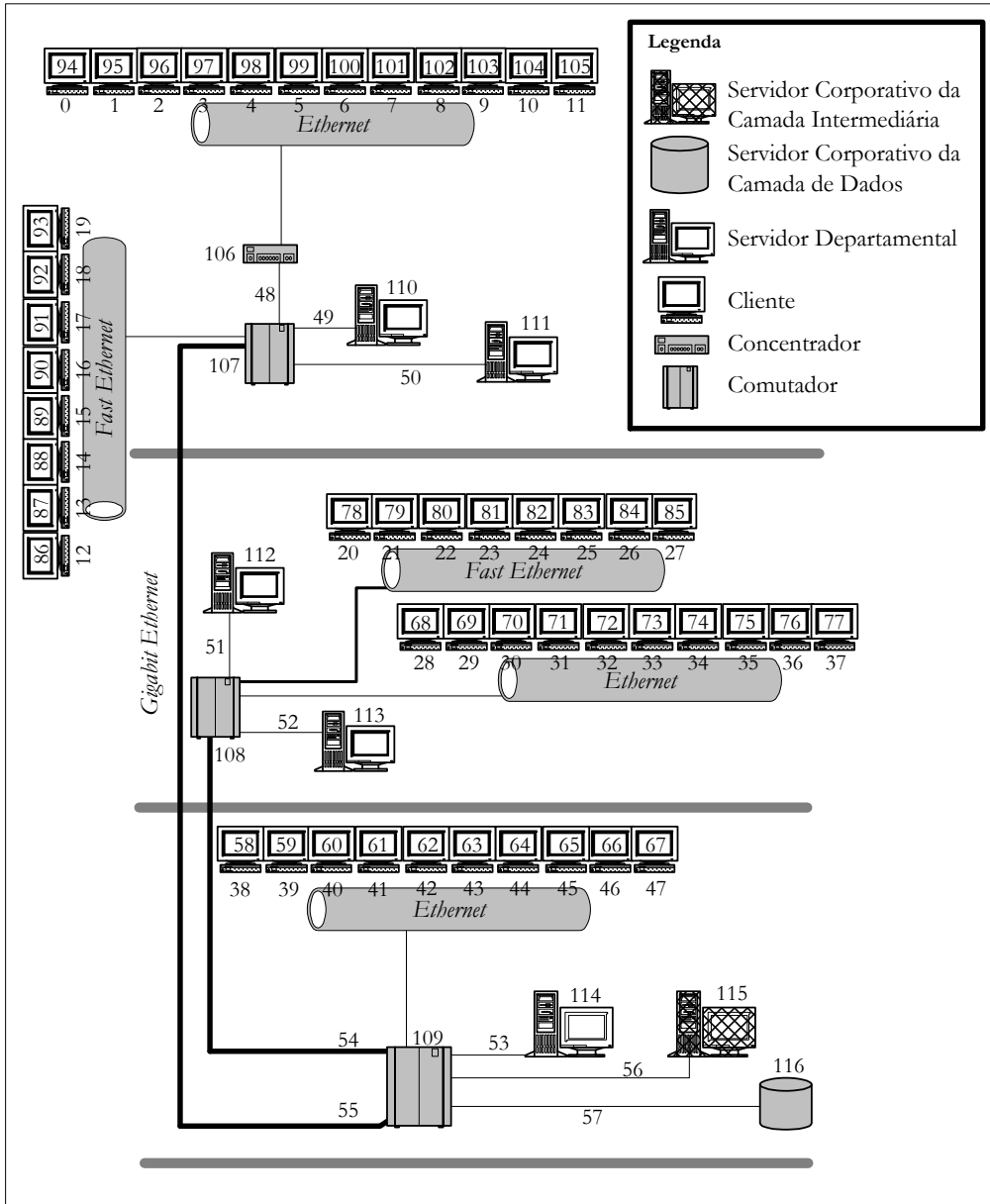


Figura C.11 - Configuração Topológica 11

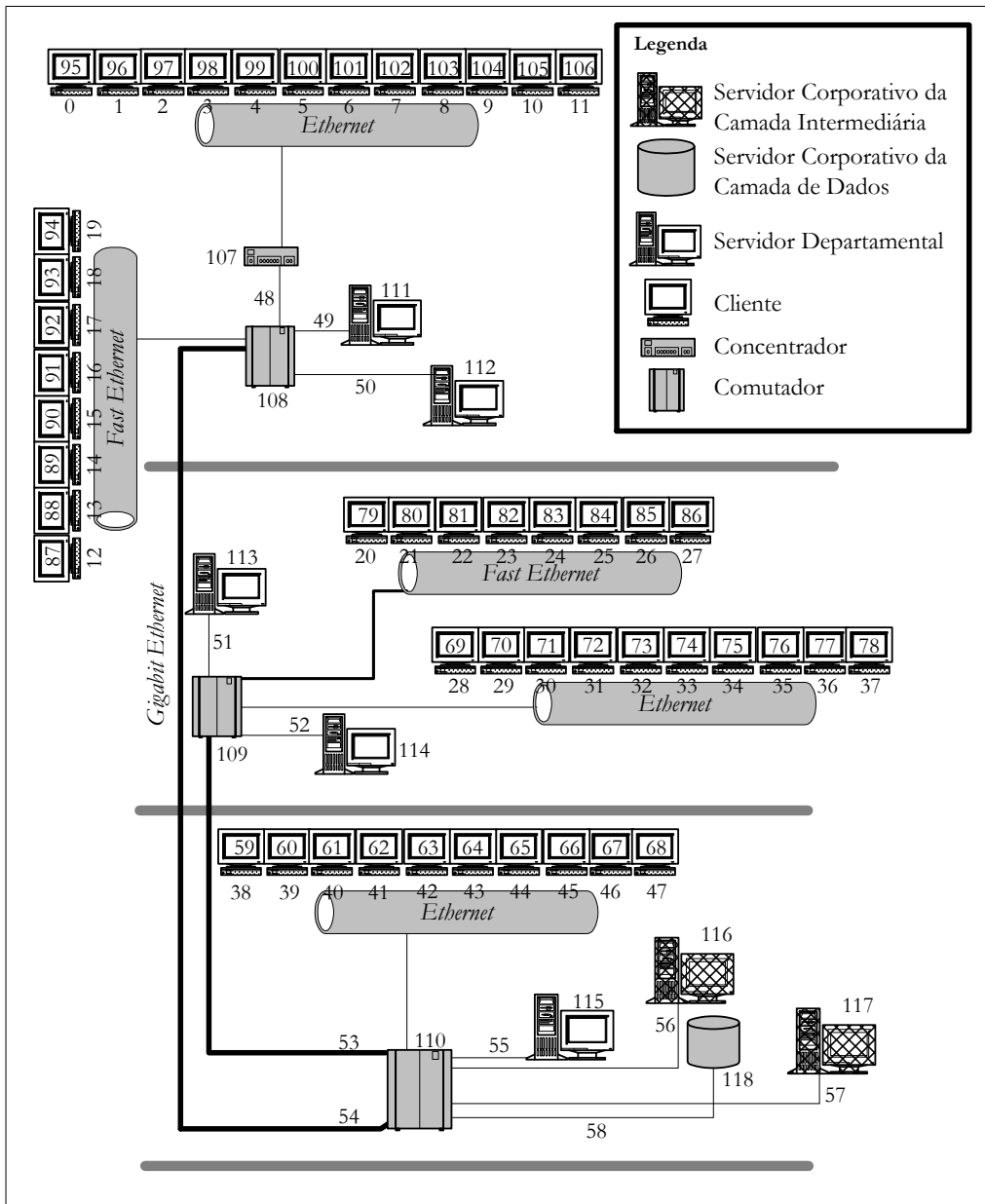


Figura C.12 - Configuração Topológica 12

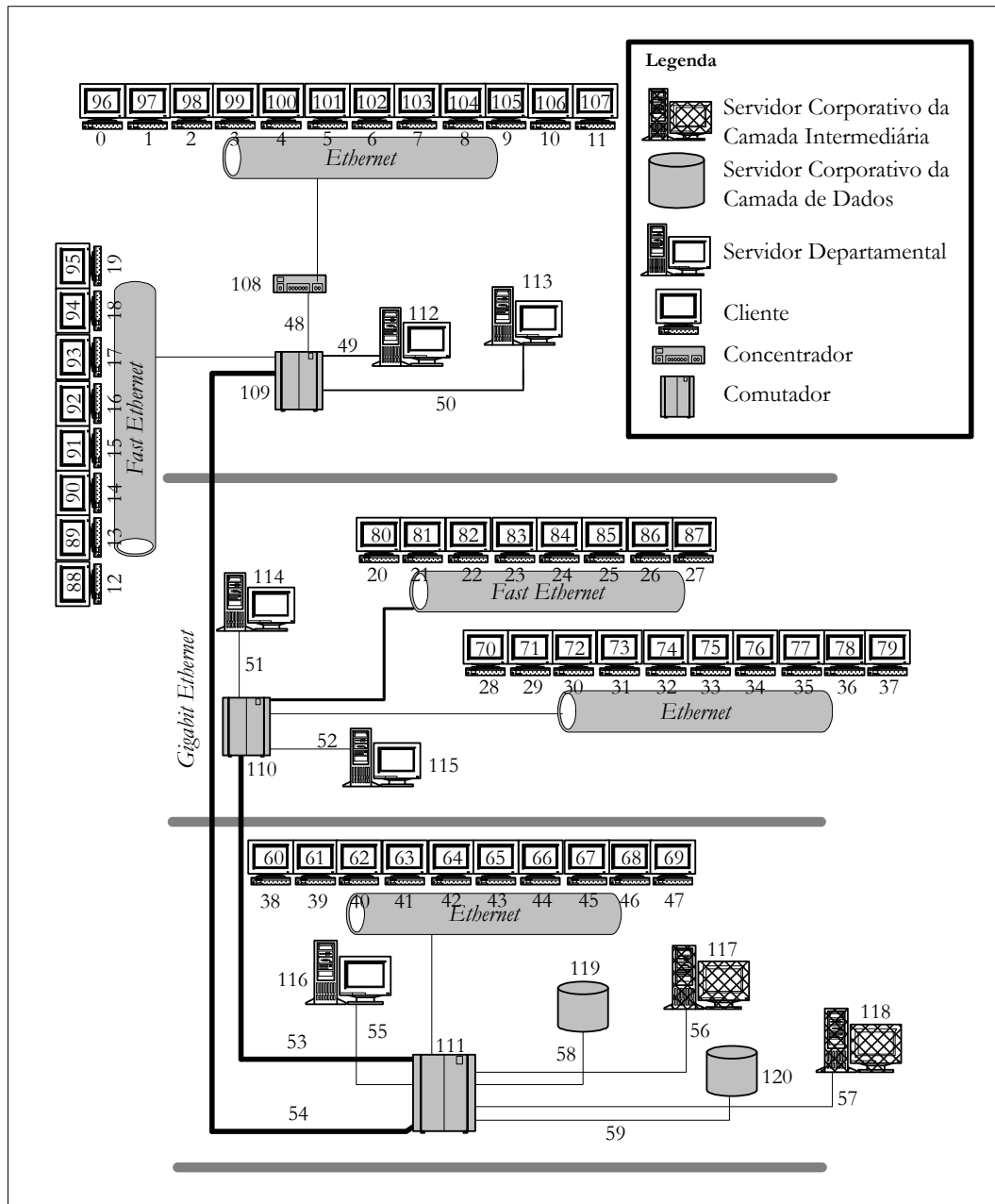


Figura C.13 - Configuração Topológica 13

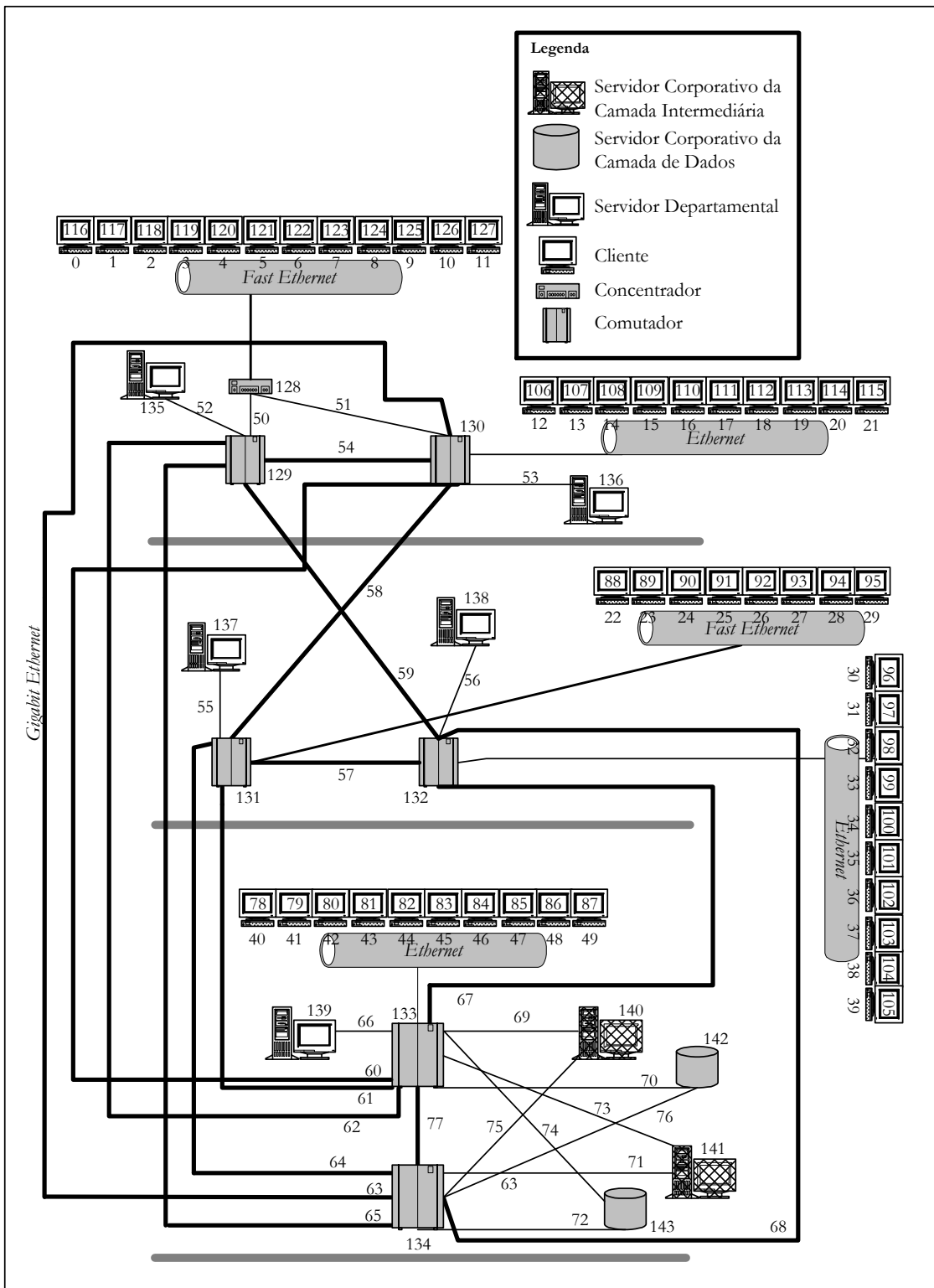


Figura C.14 - Configuração Topológica 14

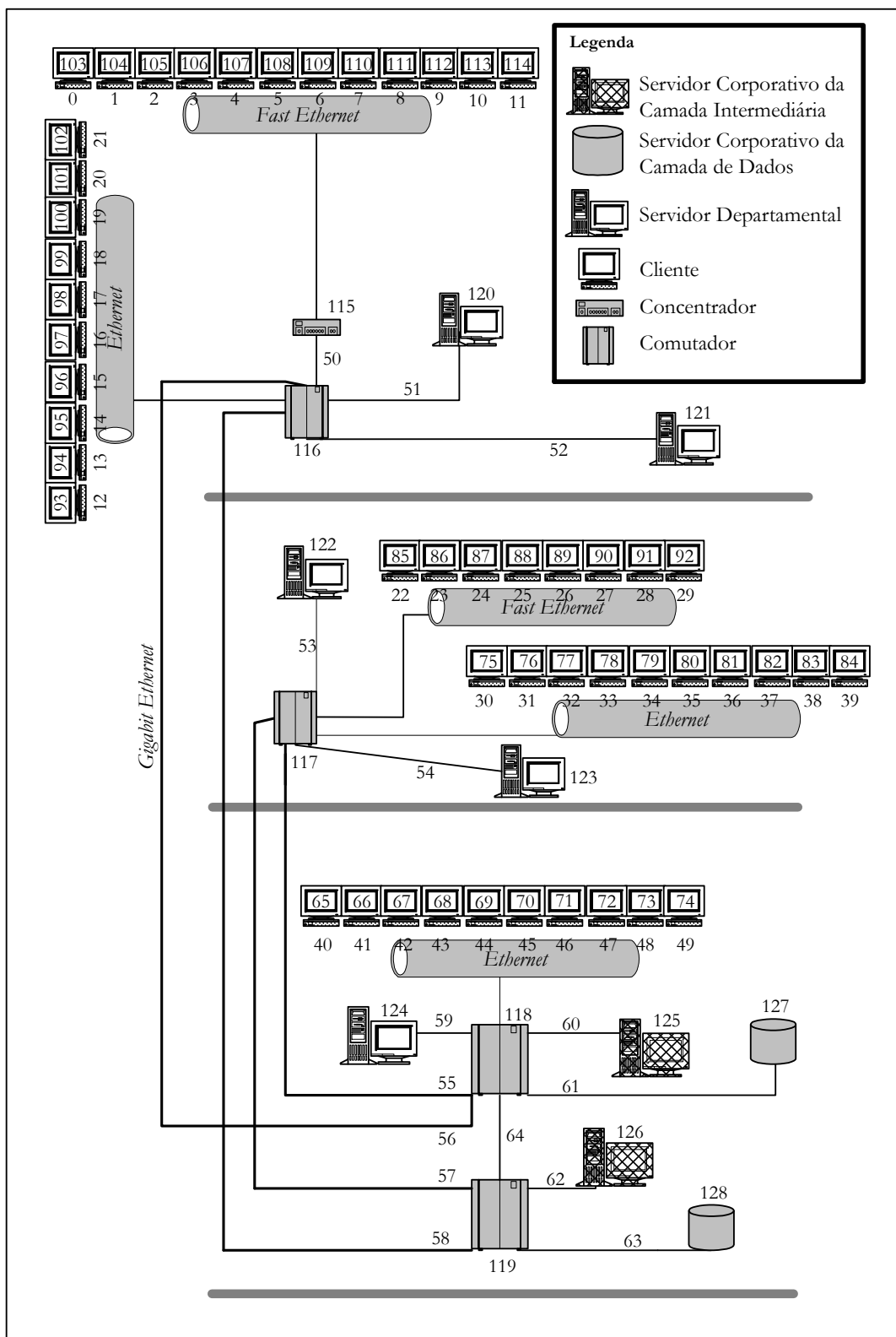


Figura 15 - Configuração Topológica 15

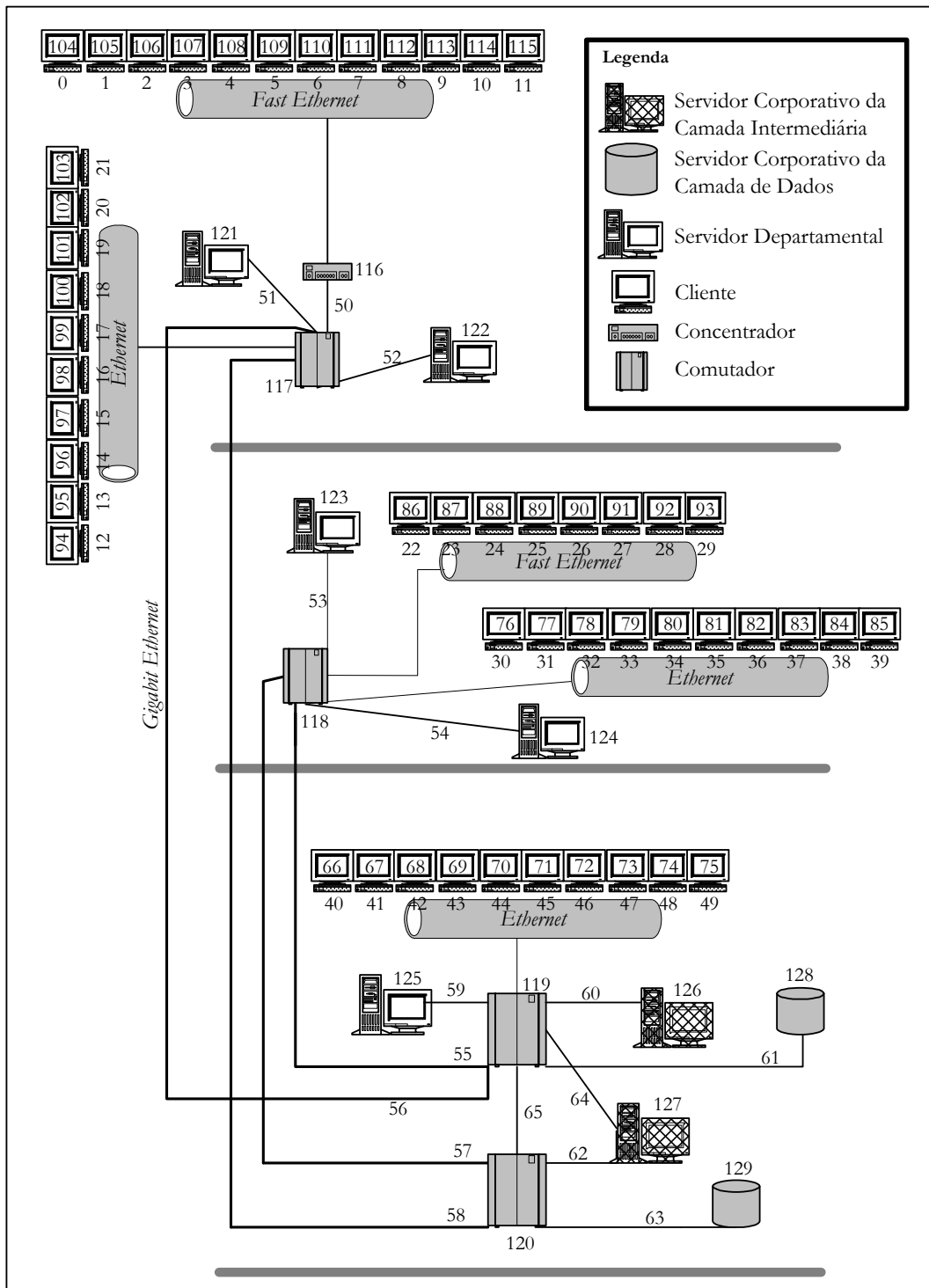


Figura C.16 - Configuração Topológica 16

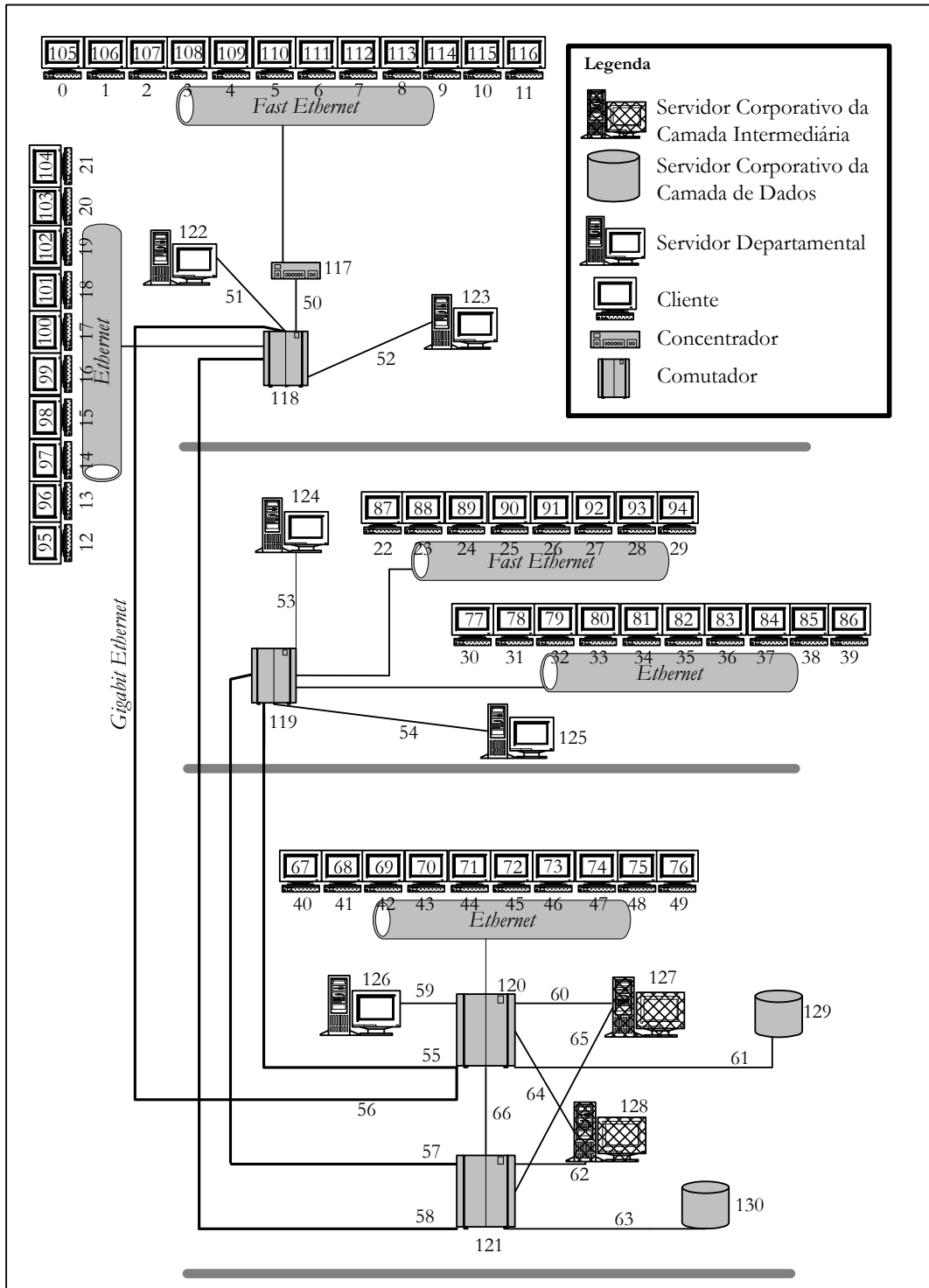


Figura C.17 - Configuração Topológica 17

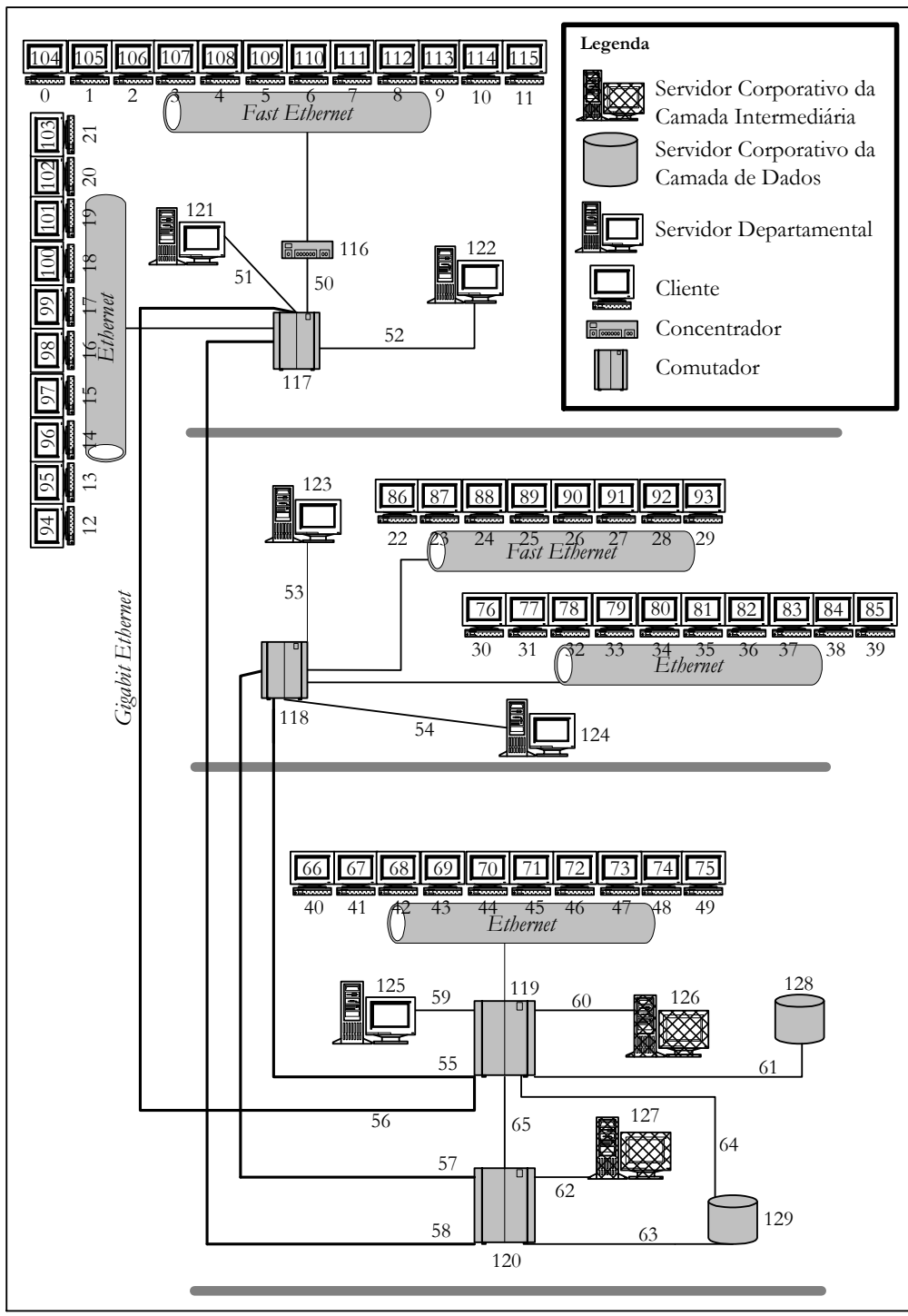


Figura C.18 - Configuração Topológica 18

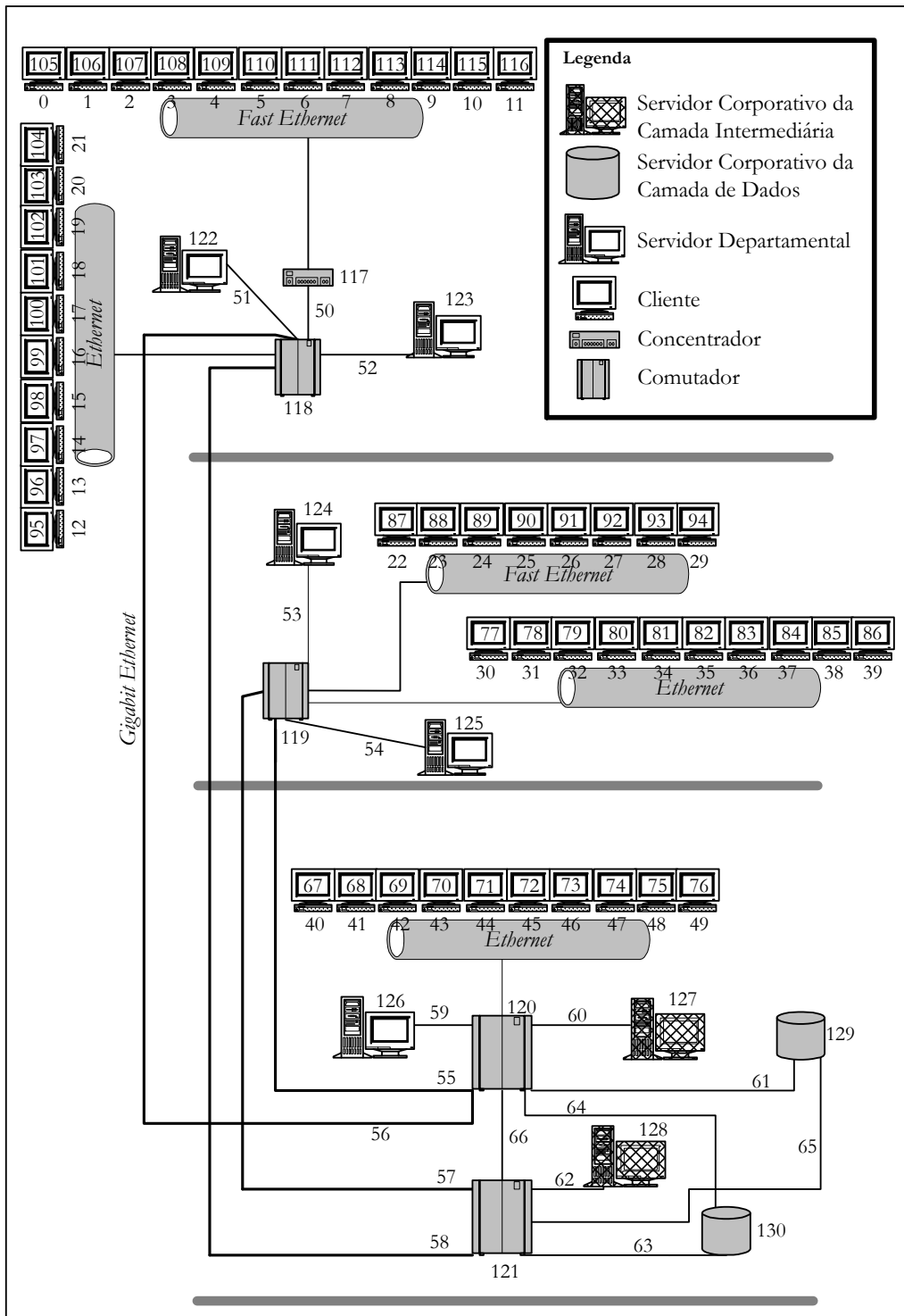


Figura C.19 - Configuração Topológica 19

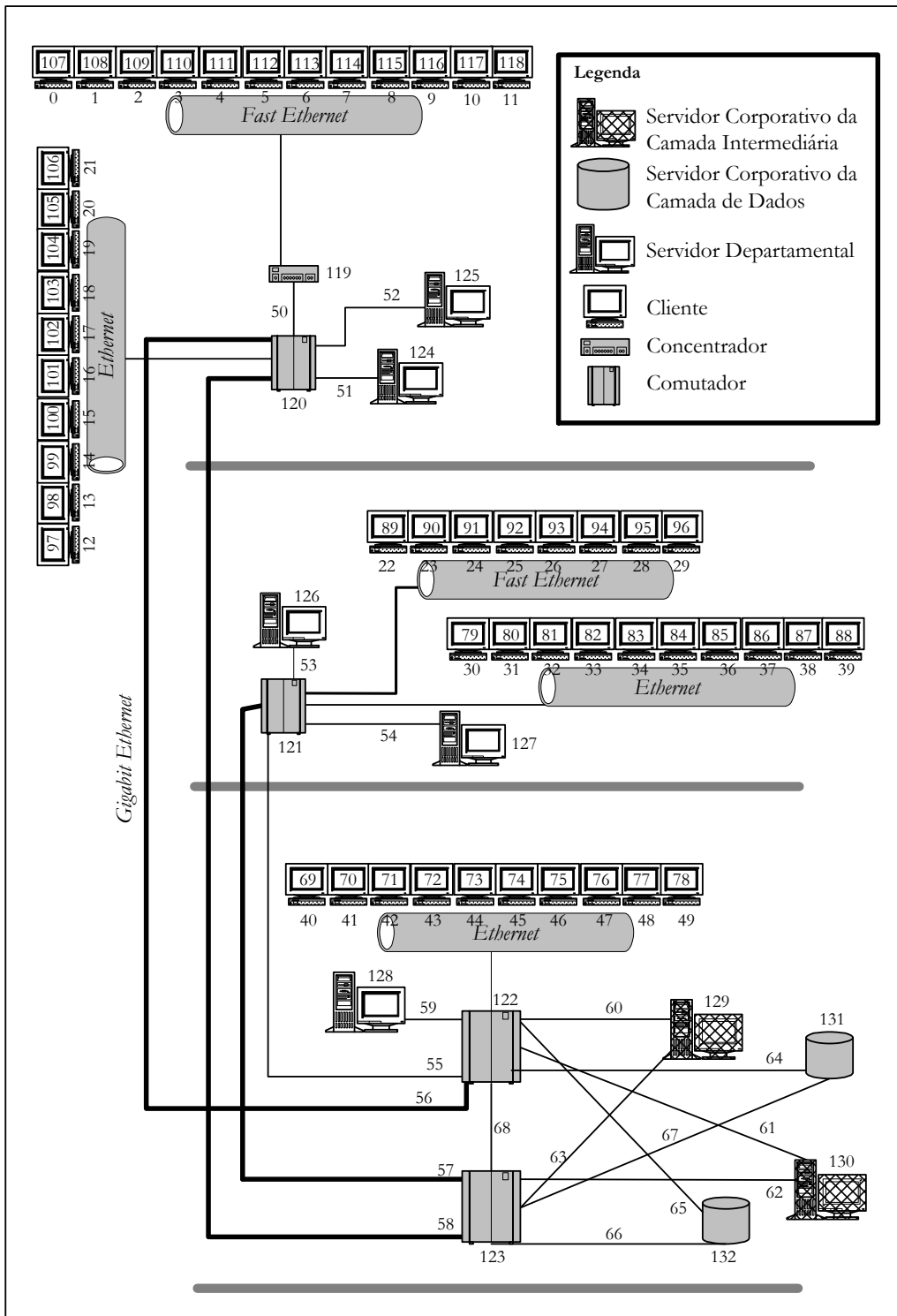


Figura C.20 - Configuração Topológica 20

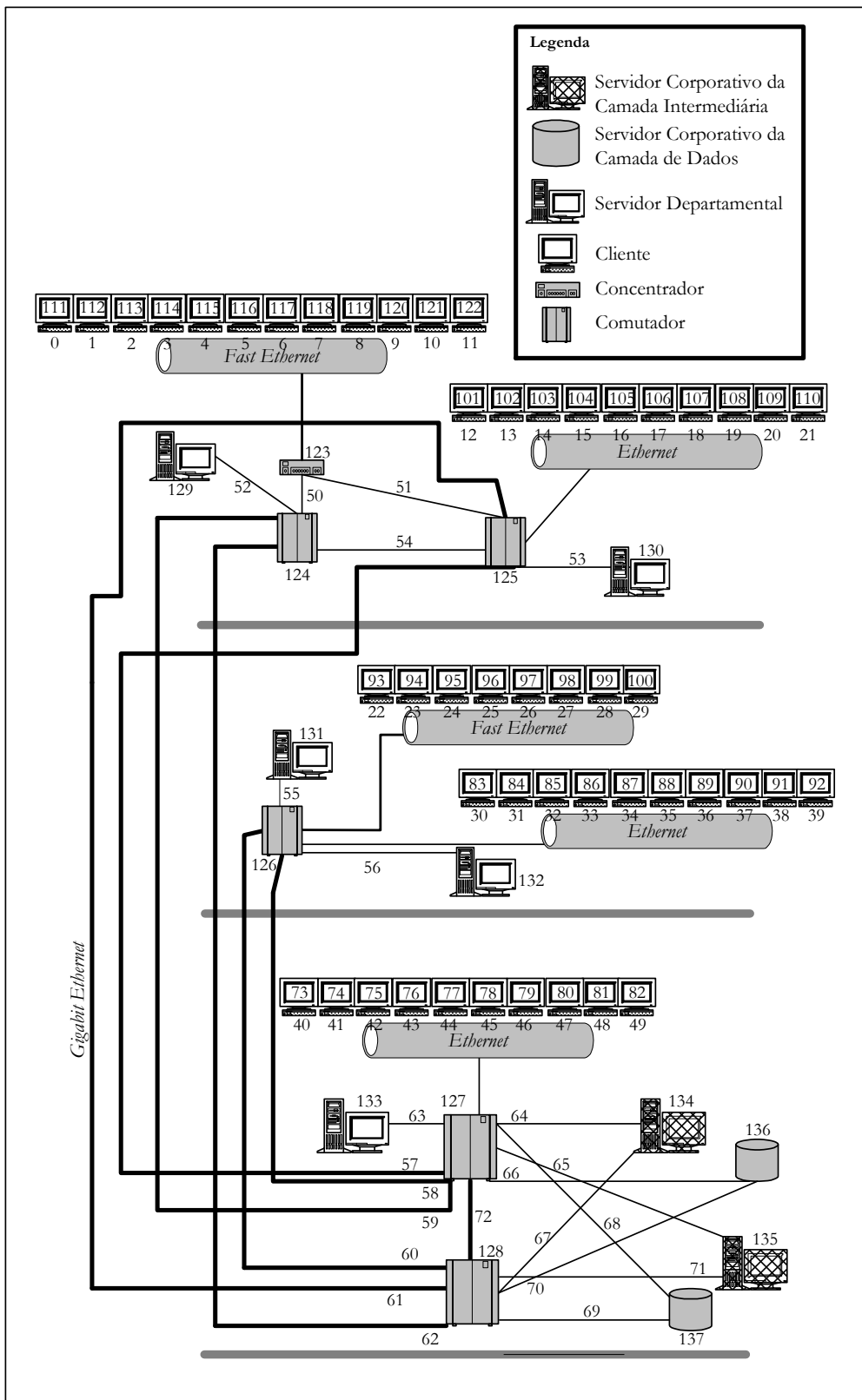


Figura C.21 - Configuração Topológica 21

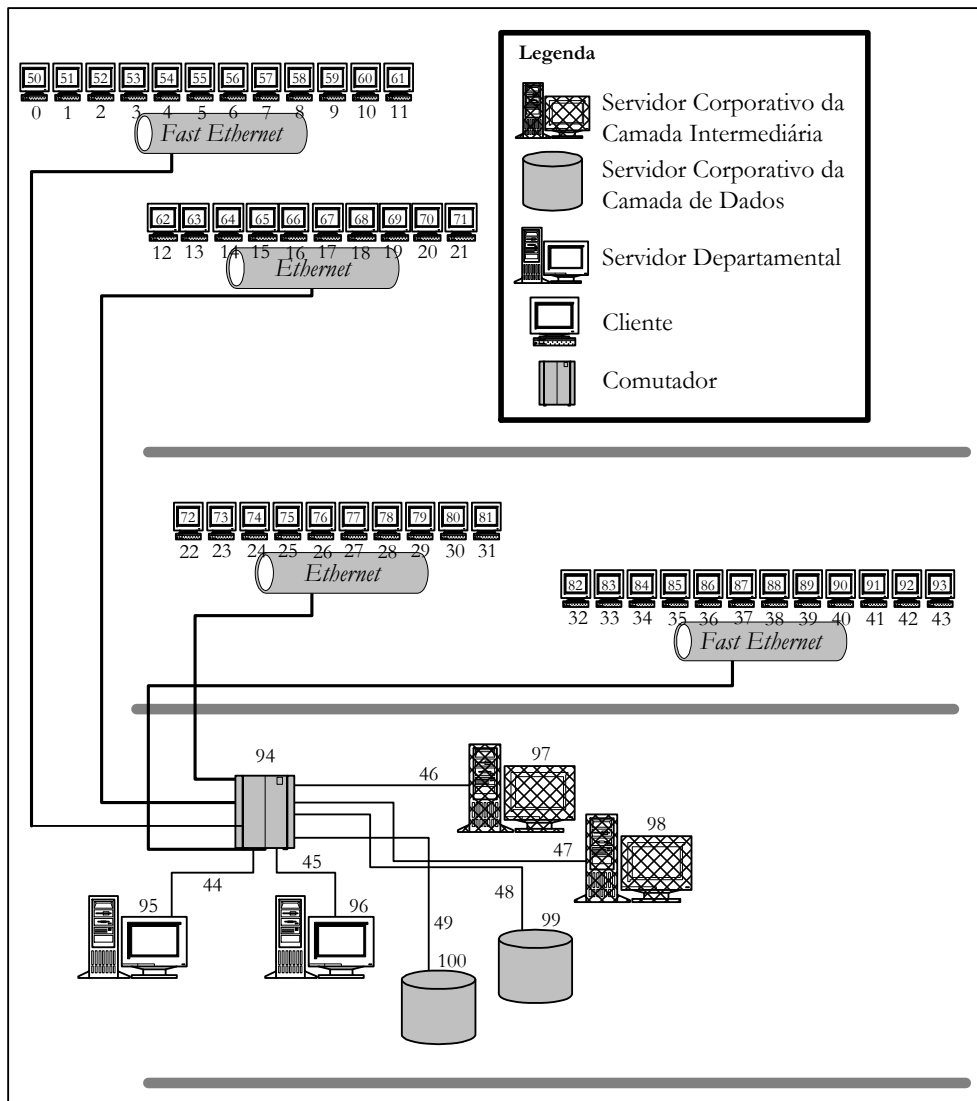


Figura C.22 - Configuração Topológica 22

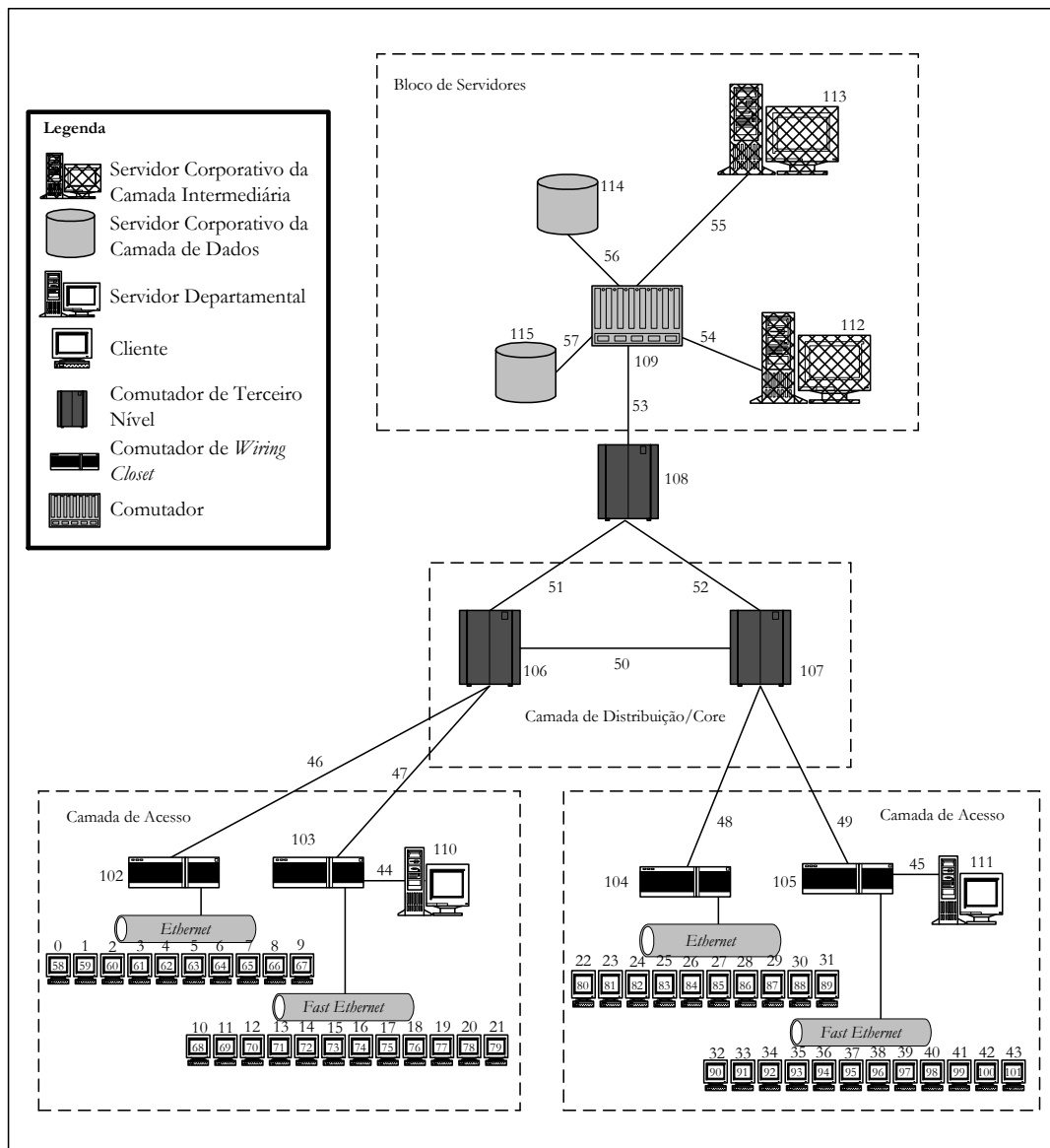


Figura C.23 - Configuração Topológica 23