Universidade de Aveiro Departamento de Electrónica e Telecomunicações,
2001

Amaro Fernandes de Sousa

Dimensionamento de redes ATM multi-serviço: modelos de programação linear inteira

Tese apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Engenharia Electrotécnica, realizada sob a orientação científica do Prof. Dr. Oliveira Duarte, Professor Catedrático do Departamento de Electrónica e Telecomunicações da Universidade de Aveiro e do Prof. Dr. Rui Valadas, Professor Associado do Departamento de Electrónica e Telecomunicações da Universidade de Aveiro

o júri

presidente

Prof. Doutor Dinis Gomes de Magalhães dos Santos
professor catedrático da Universidade de Aveiro por delegação do Reitor da Universidade de Aveiro

Prof. Doutor Aníbal Manuel de Oliveira Duarte
professor catedrático da Universidade de Aveiro (Orientador)

Prof. Doutor Carlos Henggler Antunes
professor associado da Faculdade de Ciências e Tecnologia da Universidade de Coimbra

Prof. Doutor Luís Eduardo Neves Gouveia
professor associado da Faculdade de Ciências da Universidade de Lisboa

Prof. Doutor Rui Jorge Morais Tomás Valadas
professor associado da Universidade de Aveiro (Orientador)

Prof. Doutor Nelson Fernando Pacheco da Rocha
professor associado da Universidade de Aveiro

Prof. Doutor João Luís da Costa Campos Gonçalves Sobrinho
professor auxiliar do Instituto Superior Técnico da Universidade Técnica de Lisboa

agradecimentos

Ao Prof. Oliveira Duarte cujo convite para fazer o Doutoramento sob a sua orientação foi a razão principal da concretização do trabalho apresentado nesta tese.

Ao Prof. Rui Valadas pelo acompanhamento, pelo rigor, pela crítica permanente, pelas estimulantes sessões de trabalho conjunto e pela compreensão manifestada.

Ao Prof. Luís Gouveia pelas sessões de trabalho conjunto que me permitiram encarar de uma forma segura os aspectos de Investigação Operacional necessários à concretização deste trabalho.

Ao Engº Luís Cardoso, Engº Carlos Lopes e Engº Ricardo Cadime cuja ajuda no desenvolvimento de partes significativas do software necessário ao trabalho apresentado foi de extrema importância para a sua concretização.

Ao Departamento de Electrónica e Telecomunicações e ao Instituto de Telecomunicações, pólo de Aveiro, pelas condições que me permitiram concretizar este trabalho.

À Fundação para a Ciência e Tecnologia (anteriormente designada por JNICT) pela bolsa de investigação científica do programa PRAXIS que usufruí na fase inicial do meu trabalho.

Ao Fundo Social Europeu pelo apoio dado à Universidade de Aveiro através do programa PRODEP que me proporcionou melhores condições de trabalho.

À Paula e ao Diogo cujo carinho e paciência sempre serviram de estímulo, principalmente nas fases mais difíceis do trabalho.

resumo

Esta tese aborda o problema do dimensionamento de redes ATM (*Asynchronous Transfer Mode*) num cenário multi-serviço de suporte a serviços com diferentes características de tráfego: serviços conversacionais, serviços de consulta e serviços ponto-multiponto. O problema estudado considera a perspectiva de um operador de telecomunicações que tem como objectivo a determinação da rede física com o menor custo de operação e manutenção possível e que seja capaz de suportar o tráfego esperado para os diferentes serviços com a qualidade de serviço necessária. Os problemas de dimensionamento abordados consideram serviços comutados em que cada serviço é suportado por uma rede lógica constituída por VPCs (*Virtual Path Connections*) do tipo DBR (*Deterministic Bit Rate*) e o encaminhamento é fixo (uni-horário e multi-horário).

O primeiro problema endereçado é o dimensionamento de redes ATM em que as redes lógicas de suporte a todos os tipos de serviços são constituídas exclusivamente por VPCs extremo-a-extremo. Esta arquitectura tem como principal vantagem a simplificação dos mecanismos de CAC (*Call Admission Control*) da rede. São propostos modelos de Programação Inteira para este problema. Os modelos desenvolvidos extendem modelos previamente conhecidos para serviços conversacionais de modo a poder ser considerado tráfego assimétrico (característico dos serviços de consulta e serviços ponto-multiponto) e a modelização de VPCs em que um dos extremos não é conhecido à partida (característico dos serviços de consulta quando existem múltiplos servidores na rede). É proposto um algoritmo heurístico baseado na técnica da Relaxação Lagrangeana com Optimização por Sub-Gradiente para a resolução dos modelos desenvolvidos em redes de grande dimensão.

No entanto, uma arquitectura baseada em VPCs extremo-a-extremo conduz a uma utilização pouco eficiente dos recursos da rede e, para redes de maiores dimensões, torna-se impraticável pelo elevado número de VPCs que exige. Como alternativa, são propostas duas arquitecturas diferentes. Para serviços conversacionais, é proposta uma arquitectura hierarquizada a 2 níveis para as redes lógicas de suporte. Esta arquitectura permite reduzir significativamente o número total de VPCs que é necessário configurar na rede sem contudo permitir que o número máximo de VPCs atravessados por cada ligação seja maior que três (um factor determinante para serviços comutados em que o tempo de estabelecimento de chamada é um parâmetro de qualidade de serviço a ter em consideração). Para serviços ponto-multiponto, é proposta uma arquitectura em árvore para as redes lógicas de suporte. São propostos modelos de Programação Inteira e Programação Inteira Mista para os problemas de dimensionamento da rede com as redes lógicas segundo as arquitecturas propostas. A resolução dos modelos propostos através de um algoritmo de *branch-and-bound* permite demonstrar que as arquitecturas alternativas propostas conduzem a ganhos significativos nos custos das redes obtidas quando comparados com os custos que se obtêm com redes lógicas baseadas em VPCs extremo-a-extremo.

abstract

This thesis addresses the problem of dimensioning ATM networks in a multi-service scenario for the simultaneous support of different types of services: conversational services, retrieval services and point-multipoint services. We consider the problem in a network operator's point of view, which aims to find the least operational & maintenance cost physical network that can support the expected traffic with the required quality of service. All the dimensioning problems addressed in this work consider switched services where each service is supported by a logical network. Each logical network is composed by a set of constant bit rate Virtual Path Connections (VPCs) with static routing (uni-hour and multi-hour).

The first problem addressed in this thesis is the dimensioning of ATM networks where all kinds of services are supported by logical networks based on end-to-end VPCs. This architecture greatly simplifies Call Admission Control (CAC) procedures. We propose appropriate Integer Programming models for the problem. The proposed models extend previously known models for conversational services in order to cope with assimetrical traffic (required by retrieval and point-multipoint services) and VPC modeling where one of its end nodes is not known in advance (required by retrieval services when there is more than one server node). We propose an heuristic algorithm based on Lagrangean Relaxation with Sub-Gradient Optimization for the resolution of the proposed models in large networks.

However, this architecture leads to low efficiency use of network resources. Moreover, the required large number of VPCs makes it infeasible for large networks. As an alternative, two different logical network architectures are proposed. In the case of conversational services, a 2-layer hierarchical architecture is proposed for the supporting logical networks. This architecture reduces the required number of VPCs and, simultaneously, limits to three the maximum number of VPCs crossed by each service call (an important issue for switched services where call set-up time is a quality of service parameter that must be taken into account). In the case of point-multipoint services, a tree architecture is proposed for the supporting logical networks. Different Integer Programming and Mixed Integer Programming models are proposed for the network dimensioning problem with logical networks following the proposed architectures. The use of a branch-and-bound algorithm to solve the proposed models has demonstrated that the alternative architectures achieve significant gains in the cost of the dimensioned networks when compared with the costs obtained using logical neworks based on end-to-end VPCs.

índice

Capítulo 1 Introdução....1
1.1 Tecnologia ATM....2
1.1.1 Multiplexagem Assíncrona da Informação....3
1.1.2 Tecnologia Orientada à Ligação....4
1.1.3 Redes Lógicas baseadas em VPCs....6
1.1.4 Gestão de Recursos e Dimensionamento....8
1.2 Tema e Objectivos Principais....8
1.3 Evolução do Estado-da-Arte....12
1.4 Estrutura da Tese....16
1.5 Principais Contribuições....18

Capítulo 2 Conceitos Base de Dimensionamento....21
2.1 Modelo da Rede Física....22
2.2 Modelo do Custo da Rede Física....24
2.2.1 Custos de Transmissão....24
2.2.2 Custos de Comutação....25
2.2.3 Modelo de Custos Adoptado....26
2.3 Modelo de uma Rede Lógica....26
2.3.1 Modelização Explícita de VPCs....27
2.3.2 Modelização de VPCs por Trajectos Candidatos....28
2.4 Modelo de Tráfego Oferecido....30
2.4.1 Nível da Ligação de Tráfego Ponto-a-Ponto....31
2.4.2 Nível da Ligação de Tráfego Ponto-Multiponto....33
2.4.3 Nível da Célula....38
2.5 Definição de Probabilidade de Bloqueio....39
2.6 Problema de Dimensionamento: Integralidade dos VPCs....41

Capítulo 3 Redes Lógicas com VPCs Extremo-a-Extremo....45
3.1 Cálculo da Capacidade dos VPCs....46
3.1.1 Serviços Conversacionais....46

3.1.2 Serviços de Consulta....47
3.1.3 Serviços Ponto-Multiponto....48
3.2 Modelo de Dimensionamento Uni-Horário....49
3.2.1 Formulação por Modelização Explícita de VPCs....51
3.2.2 Formulação por Trajectos Candidatos....52
3.3 Modelo de Dimensionamento Multi-Horário....53
3.3.1 Formulação por Modelização Explícita de VPCs....54
3.3.2 Formulação por Trajectos Candidatos....55
3.4 Resolução dos Modelos de Dimensionamento....57
3.4.1 Relaxação Lagrangeana com Optimização por Sub-Gradiente....57
3.4.2 Algoritmo Implementado....60
3.5 Resultados Computacionais....65
3.5.1 Definição dos Problemas Considerados....66
3.5.2 Modelização Explícita vs. Modelização por Trajectos Condidatos....69
3.5.3 Dimensionamento Uni-Horário vs. Multi-Horário....72
3.5.4 Análise do Desempenho do Algoritmo....73
3.5.5 Análise do Limite Inferior do Problema....81
3.5.6 Observações Finais....84
3.6 Conclusões....84

Capítulo 4 Redes Lógicas Hierarquizadas a Dois Níveis....87
4.1 Arquitectura das Redes Lógicas....88
4.2 Aproximações Consideradas....90
4.3 Cálculo das Capacidades dos VPCs....93
4.4 Modelo de Dimensionamento Uni-Horário....96
4.5 Modelo de Dimensionamento Multi-Horário....99
4.6 Resultados Computacionais....102
4.7 Conclusões....105

Capítulo 5 Redes Lógicas em Árvore....107
5.1 Abordagem ao Problema....109
5.1.1 Majorante da Probabilidade de Bloqueio....110
5.1.2 Modelos Lineares da Função ErlangB Inversa....114
5.1.3 Modelo Binário da Função ErlangB Inversa....116
5.1.4 Modelo Real da Função ErlangB Inversa....119
5.2 Modelos de Dimensionamento Uni-Horário....122

5.2.1 Modelo A de Dimensionamento Uni-Horário 126
5.2.2 Modelo B de Dimensionamento Uni-Horário 129
5.3 Reformulações dos Modelos de Dimensionamento Uni-Horário 131
5.3.1 Modelo A Reformulado de Dimensionamento Uni-Horário 133
5.3.2 Modelo B Reformulado de Dimensionamento Uni-Horário 136
5.4 Modelos de Dimensionamento Multi-Horário 138
5.4.1 Modelo A de Dimensionamento Multi-Horário 138
5.4.2 Modelo B de Dimensionamento Multi-Horário 140
5.4.3 Modelo A Reformulado de Dimensionamento Multi-Horário 143
5.4.4 Modelo B Reformulado de Dimensionamento Multi-Horário 145
5.5 Resultados Computacionais 148
5.6 Conclusões 157

Capítulo 6 Considerações Finais 159
6.1 Principais Conclusões 159
6.2 Sugestões para Trabalho Futuro 163

Anexo A Resolução de Equações (Tráfego Ponto-Multiponto) 165

Anexo B Grafos e Resultados do Capítulo 3 169

Anexo C Algoritmo *Branch-and-Bound* 185

Lista de Acrónimos 191

Referências 193

CAPÍTULO 1

INTRODUÇÃO

O domínio das redes de telecomunicações sofreu importantes transformações ao longo das últimas décadas. Dois dos mais importantes factores nestas transformações são o aparecimento de tecnologias de transmissão mais avançadas e a identificação de potenciais novos serviços. Por um lado, os avanços da micro-electrónica e da fotónica permitiram a implementação de sistemas de transmissão da ordem dos gigabits por segundo. Por outro lado, o crescimento constante da capacidade de processamento e armazenamento de informação dos equipamentos terminais e o desenvolvimento de métodos mais eficientes de processamento de som, imagem e vídeo conduziram ao aparecimento de novas aplicações telemáticas com requisitos mais exigentes de comunicação tais como maior capacidade e flexibilidade de transferência de informação. Algumas décadas atrás, as redes de telecomunicações públicas de cada país eram geridas por operadores aos quais, pela sua posição de monopólio, era possível fazer investimentos com ciclos de retorno de investimento muito grandes. Nas últimas décadas, no entanto, a abertura do mercado das telecomunicações a diferentes operadores dificultou a capacidade de resposta dos operadores na implementação de redes que pudessem tirar partido das novas tecnologias de transmissão de modo a oferecer os novos serviços identificados. Num mercado de livre concorrência, não é economicamente viável um operador desenvolver e pôr em operação

diferentes redes de telecomunicações para suportar diferentes serviços, cada um sobre a tecnologia de transmissão mais adequada. A alternativa era o desenvolvimento de uma rede de alto débito que permitisse o suporte simultâneo de múltiplos serviços. No entanto, esta tarefa era extremamente difícil de ser levada a cabo individualmente por cada operador de telecomunicações. Estas condicionantes conduziram a que os fabricantes de equipamentos de telecomunicações e os operadores de telecomunicações cooperassem entre si de modo a especificarem normas que permitissem a implementação de uma Rede Digital com Integração de Serviços de Banda Larga (RDIS-BL). Como resultado deste esforço surgiu a norma ATM (*Asynchronous Transfer Mode*) [CCITT92a] [CCITT92b] que define um protocolo de transporte capaz de, por um lado, ser implementado sobre qualquer tecnologia de transmissão e, por outro lado, suportar diferentes tipos de serviços.

1.1 Tecnologia ATM

Uma rede ATM utiliza um mecanismo de comutação rápida de pacotes (tradução do termo anglo-saxónico *fast packet switching*) em que a integração dos serviços é conseguida pela segmentação de todas as fontes de informação em pacotes de tamanho único designados por células ATM. Por outro lado, uma rede ATM é orientada à ligação, ou seja, para que um equipamento terminal troque informação com outro, é necessário que uma ligação virtual esteja estabelecida entre ambos. A fase de estabelecimento de uma ligação virtual define um trajecto na rede que será percorrido por todas as células durante o tempo de vida da ligação. No caso geral, uma rede ATM é constituída por um conjunto de comutadores ATM ligados entre si por ligações ponto-a-ponto em que cada ligação física tem uma determinada capacidade de transmissão. A norma não impõe qualquer limitação topológica para a interligação dos comutadores nem restringe o tipo de ligações possíveis entre eles. Assim, é possível a implementação de ligações físicas entre comutadores ATM usando qualquer outra tecnologia de transmissão disponível. Estas ligações podem ser simples ligações físicas ponto-a-ponto (um par de fibras ópticas a ligar directamente dois comutadores) ou tributários de uma qualquer rede de comutação tal como PDH (*Plesiochronous Digital Hierarchy*) ou SDH (*Synchronous Digital Hierarchy*).

1.1.1 Multiplexagem Assíncrona da Informação

Numa rede ATM, a integração de diferentes serviços é conseguida pela segmentação de todas as fontes de informação em pacotes de formato único designados por células ATM. Esta integração é feita por multiplexagem assíncrona das células de diferentes fontes de informação em cada uma das ligações físicas da rede. A Figura 1-1 ilustra este conceito. O termo "assíncrono" é usado para qualificar a multiplexagem de diferentes fontes de informação reflectindo o facto das células de cada fonte poderem ser transmitidas em intervalos de tempo irregulares. Na ligação física circulam continuamente células ao ritmo máximo da capacidade da ligação e são enviadas células vazias sempre que à entrada da ligação não existirem células de informação.

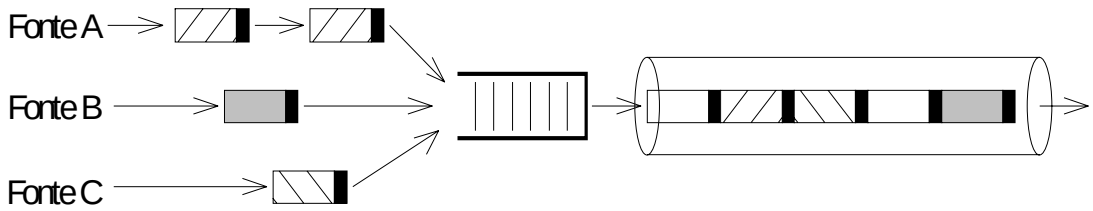


Figura 1-1 : Multiplexagem de diferentes fontes de informação numa ligação física ATM

O modo de transferência ATM, conforme ilustrado na Figura 1-1, tem três características fundamentais que o distinguem das tecnologias tradicionais de comutação de circuitos. A primeira característica é o facto do ritmo máximo de transmissão de cada fonte estar apenas limitado pela capacidade da ligação física (também designada por largura de banda). Esta característica permite a integração de fontes de informação com ritmos de transmissão diferentes na mesma rede física. Outra característica importante é o facto de, quando uma fonte não está a transmitir células, as outras fontes poderem usar a capacidade da ligação para o fazer. Assim, uma ligação física pode suportar várias fontes de ritmo variável tal que a soma dos ritmos de pico de cada fonte seja superior à sua capacidade. Esta característica conduz a ganhos de multiplexagem estatística quando se encaminham diferentes fontes de ritmo variável pelos mesmos recursos da rede. Finalmente, o facto de em cada ligação física circularem continuamente pacotes de tamanho fixo, por um lado, torna os mecanismos de sincronismo de célula na recepção mais simples e fiáveis e, por outro lado, simplifica a arquitectura das matrizes de comutação a implementar nos nós da rede (características típicas de redes de comutação rápida de pacotes).

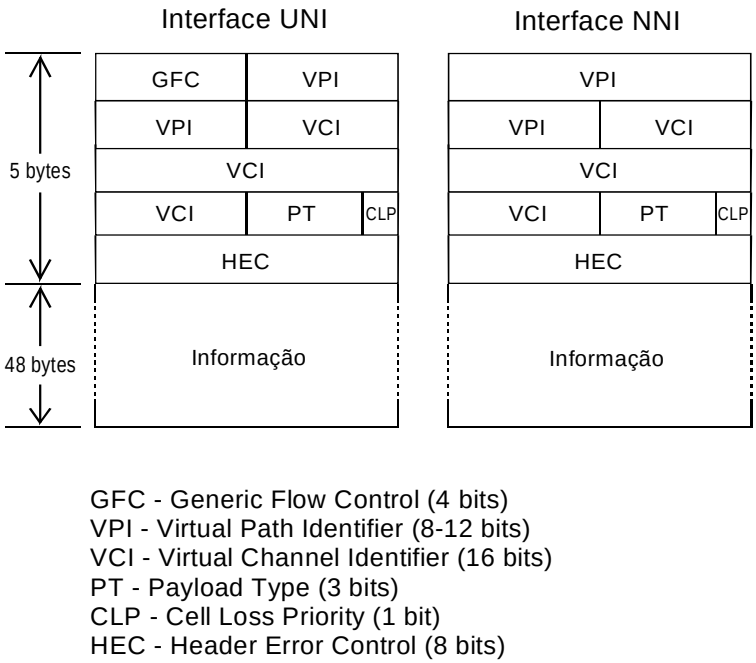


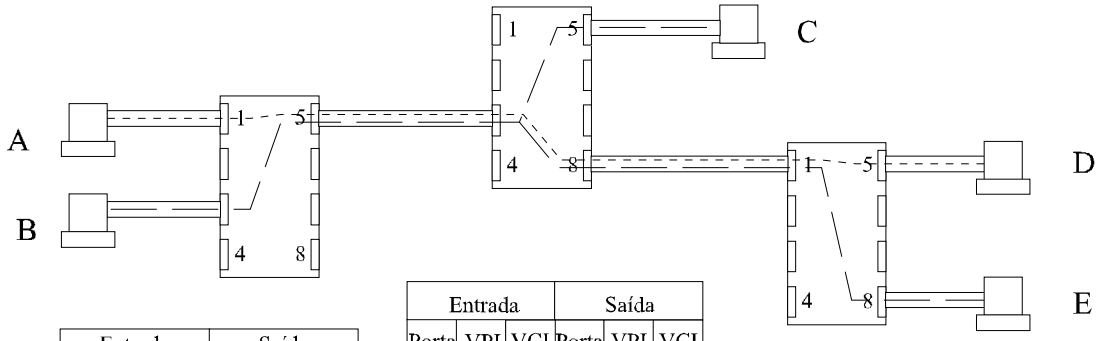
Figura 1-2 : Formatos da célula ATM

Uma célula ATM tem um tamanho fixo de 53 bytes e é constituída por um cabeçalho de 5 bytes e por um campo de informação de 48 bytes. Distinguem-se dois tipos de interface numa rede ATM: a interface UNI (*User-Network Interface*) entre um equipamento terminal e um comutador ATM e a interface NNI (*Network-Network Interface*) entre quaisquer dois comutadores ATM. Os campos do cabeçalho de uma célula ATM são ligeiramente diferentes para os dois tipos de interface (Figura 1-2). Dos vários campos que constituem o cabeçalho, os que são relevantes para o tema desta tese são o identificador de ligação virtual (VCI – *Virtual Channel Identifier*) e o identificador de caminho virtual (VPI – *Virtual Path Identifier*). É com base nestes identificadores que em cada ligação física (Figura 1-1), o receptor sabe a que ligação lógica cada célula pertence. Pelo formato da célula ATM (interface NNI da Figura 1-2), verifica-se que em cada ligação física entre 2 comutadores ATM, o número máximo de ligações virtuais é 2^{16} = 65536 e o número máximo de caminhos virtuais é 2^{12} = 4096.

1.1.2 Tecnologia Orientada à Ligação

Uma rede ATM é orientada à ligação, ou seja, para que um equipamento terminal troque informação com outro é necessário que uma ligação virtual esteja estabelecida entre ambos. Uma ligação virtual é designada na terminologia da norma ATM por um VCC

(*Virtual Channel Connection*). Os identificadores VCI e VPI do cabeçalho de uma célula ATM não são universais mas válidos apenas na ligação física onde são usados. Na fase de estabelecimento de um VCC, são atribuidos identificadores VCI e VPI em todas as ligações físicas por onde a ligação virtual passa. Esta atribuição é feita nos comutadores por onde o VCC é estabelecido através de tabelas de encaminhamento. A Figura 1-3 ilustra o resultado do estabelecimento de um VCC ponto-a-ponto do terminal A para o terminal D e um VCC ponto-multiponto do terminal B para os terminais C e E numa rede constituída por 3 comutadores. Neste exemplo, o terminal A envia informação para D preenchendo todos as células com VPI=0 e VCI=21 e o terminal D sabe que todas as células que lhe chegam com VPI=0 e VCI=9 foram enviadas por A. O mesma raciocínio se aplica ao VCC ponto-multiponto. Neste último caso, o comutador central é responsável por duplicar as células vindas de B para os dois destinos da ligação. Depois dos VCCs estarem estabelecidos, a função mais importante de cada comutador é, baseado no VCI e no VPI de cada célula que chega a uma porta de entrada, encaminhá-la para a(s) respectiva(s) porta(s) de saída convertendo adequadamente o valor dos identificadores.



Entrada			Saída		
Porta	VPI	VCI	Porta	VPI	VCI
...	...	...	...	...	...
1	0	21	5	3	7
3	0	22	5	4	11
...	...	...	...	...	...

Entrada			Saída		
Porta	VPI	VCI	Porta	VPI	VCI
...	...	...	...	...	...
3	4	11	5 8	0 6	10 18
3	3	7	8	3	15
...	...	...	...	...	...

Entrada			Saída		
Porta	VPI	VCI	Porta	VPI	VCI
...	...	...	...	...	...
1	3	15	5	0	9
1	6	18	8	0	11
...	...	...	...	...	...

Figura 1-3 : Exemplo do estabelecimento de VCCs numa rede ATM

As redes ATM suportam serviços comutados e serviços permanentes. O estabelecimento das ligações virtuais (VCCs) é feito por sinalização nos serviços comutados e por gestão da rede nos serviços permanentes.

No caso de serviços comutados, o estabelecimento de um VCC é feito a pedido do terminal chamador imediatamente antes do envio da informação. Este pedido é processado por troca de mensagens de sinalização entre terminal chamador, nós de comutação da rede e terminal destinatário (ou terminais destinatários no caso de uma chamada ponto-multiponto). Todo o processo de estabelecimento dos VCCs designa-se na nomenclatura da norma ATM por CAC (*Call Admission Control*). Um pedido de estabelecimento de uma ligação virtual especifica o(s) endereço(s) destino, as características de geração de células que pretende enviar e os parâmetros de Qualidade de Serviço que pretende que a rede assegure. Com base nesta informação, o CAC é responsável por determinar um trajecto (ou árvore) com capacidade disponível suficiente para encaminhar a ligação virtual com os requisitos especificados e por garantir que todos os comutadores envolvidos preparam convenientemente as suas tabelas de encaminhamento. No tempo de vida da ligação lógica, as células submetidas à rede pelo equipamento origem são monitoradas por forma a eliminar aquelas que excedam as características de geração declaradas inicialmente. Este processo de monitorização designa-se na nomenclatura da norma ATM por UPC (*Usage Parameter Control*).

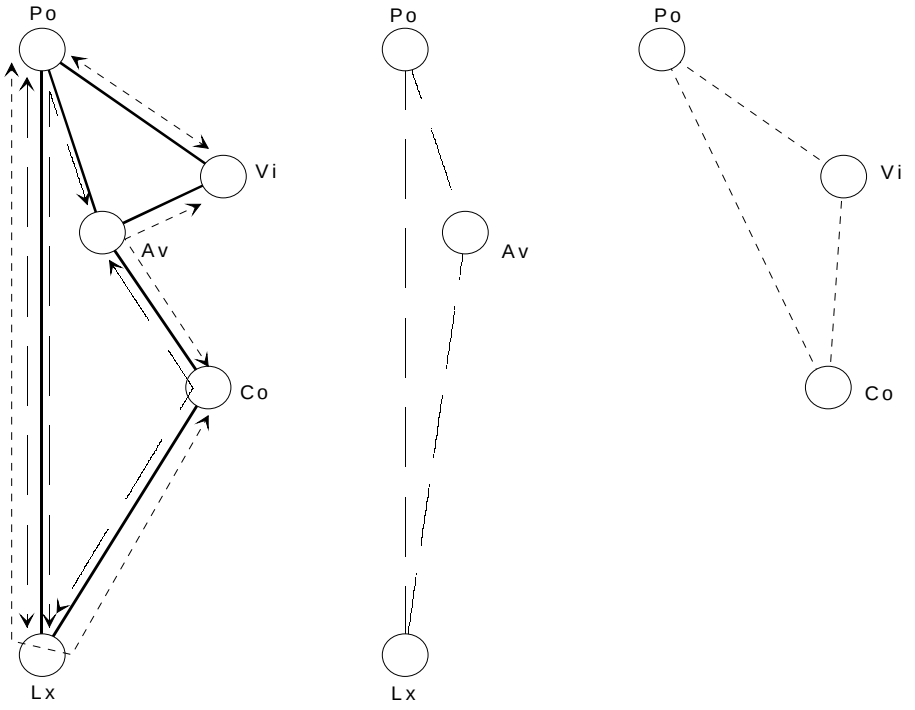
No caso de serviços permanentes, o estabelecimento dos VCCs é feito por mecanismos de gestão da rede. Neste caso, as tabelas de encaminhamento são previamente configuradas e os terminais podem enviar células de informação (com as características de geração de células previamente contratadas com o operador) sem que para isso tenham primeiro que sinalizar um pedido de ligação à rede. Neste caso, não é necessário o CAC e apenas o UPC está activo.

1.1.3 Redes Lógicas baseadas em VPCs

A tecnologia ATM permite a configuração da rede através de caminhos virtuais como elemento base para a criação de redes lógicas. Um caminho virtual é designado na terminologia da norma ATM por um VPC (*Virtual Path Connection*). Um VPC é caracterizado pelos seus pontos extremo (dois comutadores da rede), pelo trajecto físico que os liga e pela capacidade (ou largura de banda) que lhe está atribuído em todas as ligações que compõem o seu trajecto. Os VPCs são configurados por mecanismos de gestão e, por norma, são transparentes aos equipamentos terminais.

Por um VPC podem ser encaminhados diferentes VCCs em que o estabelecimento de cada um destes é decidido apenas no nó extremo do VPC. Uma vantagem óbvia do uso de

VPCs é a diminuição do tempo de estabelecimento de ligação, particularmente quando o VPC atravessa um grande número de comutadores. As experiências reportadas em [Anerousis96] demonstram que a utilização de VPCs no suporte a serviços comutados permite controlar a carga do sistema de sinalização e, assim, evitar situações da congestão da rede. Outra vantagem da utilização de VPCs é a possibilidade de configurar diferentes redes lógicas numa mesma rede física. A Figura 1-4 ilustra a criação de 2 redes lógicas sobre uma mesma rede física. A Figura 1-4a apresenta uma rede física com 5 comutadores e 6 ligações físicas entre eles na qual foram configurados 2 conjuntos de VPCs representados por dois tipos diferentes de tracejado. Os dois conjuntos de VPCs definem as duas redes lógicas representadas na Figura 1-4b e Figura 1-4c.



a) Rede Física *b) Rede Lógica 1* *c) Rede Lógica 2*

Figura 1-4 : Ilustração da utilização de VPCs na definição de redes lógicas

Diferentes serviços podem ser suportados por diferentes redes lógicas. A topologia a adoptar para cada uma das redes lógicas é, como ilustrado na Figura 1-4, independente da topologia da rede física que as suporta. No entanto, é necessário que as ligações físicas tenham uma capacidade suficiente para acomodar as capacidades de todos os VPCs que por elas passam. A forma como este cálculo é feito depende do tipo de VPCs configurados. O trabalho apresentado nesta tese considera a utilização de redes lógicas com VPCs do tipo

Deterministic Bit Rate (DBR). Neste caso, não existe multiplexagem estatística entre ligações virtuais de diferentes VPCs e a capacidade de cada ligação física da rede deve ser pelo menos igual à soma aritmética das capacidades atribuídas aos VPCs que por elas passam.

1.1.4 Gestão de Recursos e Dimensionamento

Designa-se por gestão de recursos de uma rede ATM o conjunto de procedimentos que regem o processo de estabelecimento e manutenção de VCCs na rede [Valadas98] [Dziong97] [Onvural95] [COST94]. Estes procedimentos têm em consideração as características do VCC a estabelecer e os recursos disponíveis na rede para suportar novos VCCs. A palavra "recursos" designa o conjunto de todas as capacidades de transmissão das ligações físicas entre comutadores e de todas as capacidades de encaminhamento dos comutadores existentes na rede. Quando um VCC é estabelecido, ele ocupa uma parte dos recursos em todos os elementos de rede por onde passa. A gestão de recursos determina para cada novo VCC, um trajecto (ou árvore) a adoptar no qual haja disponíveis recursos suficientes para suportar o VCC em todos os elementos de rede que o compõem.

A determinação de uma configuração física da rede capaz de suportar o tráfego previsto num determinado período de tempo futuro é designada por dimensionamento da rede. O dimensionamento de uma rede ATM deverá ter em consideração os mecanismos de gestão de recursos usados na rede durante a sua fase de operação.

A utilização de redes lógicas na configuração de redes ATM tem um papel fundamental na ligação entre os mecanismos de gestão de recursos e os processos de dimensionamento da rede. Por um lado, as redes lógicas permitem um grau de liberdade adicional na forma como a gestão de recursos pode ser implementada com o objectivo de maximizar a utilização da rede. Por outro lado, as redes lógicas permitem uma maior flexibilidade no processo de dimensionamento da rede por forma a minimizar o seu custo.

1.2 Tema e Objectivos Principais

O tema central desta tese é o estudo de metodologias adequadas ao dimensionamento de redes ATM para suporte a múltiplos serviços. Em geral, o dimensionamento da rede pressupõe a determinação da "melhor" configuração física capaz de suportar um

determinado tráfego previsto. Este objectivo é atingível em dois passos. O primeiro passo é a formulação de um modelo de optimização que defina:

(1) o conjunto das configurações físicas de rede possíveis, ou seja, as que têm recursos suficientes para suportar todo o tráfego previsto;

(2) uma medida da qualidade de cada uma das solução possíveis.

O segundo passo é a resolução do modelo de optimização através de um algoritmo que em tempo útil consiga calcular, de entre o conjunto de soluções possíveis, a solução de melhor qualidade. Quando a solução óptima não é possível ser calculada, ou porque não é conhecido nenhum algoritmo que resolva o modelo de optimização, ou porque os algoritmos conhecidos não conseguem calcular a solução óptima em tempo útil, então tenta-se determinar uma heurística (também designada por algoritmo heurístico) que calcule em tempo útil a melhor solução possível. Neste caso, devem ser determinados limites teóricos que permitam aferir da qualidade da solução encontrada. Se considerarmos por exemplo um problema de minimização, o valor de qualidade da melhor solução encontrada define um limite superior S para o valor da qualidade da solução óptima. Se for possível determinar um limite inferior teórico I do valor da qualidade da solução óptima, podemos então afirmar que a solução encontrada está no máximo a $100\cdot(S-I)/I$ por cento da solução óptima. Esta grandeza designa-se por intervalo dual (tradução do termo anglo-saxónico *duality gap*) e é a medida habitualmente usada para a determinação da qualidade das soluções calculadas por uma heurística.

Um dos objectivos do trabalho apresentado nesta tese é o de estudar a formulação dos problemas de dimensionamento de redes ATM recorrendo a modelos de programação linear. Um modelo de programação linear é um modelo de optimização formulável através de um conjunto de variáveis reais e/ou inteiras X, em que a medida de qualidade é definida por uma função linear das variáveis do problema $f(X)$ e em que o conjunto de soluções possíveis é definido por um conjunto de desigualdades da forma $g_i(X) \leq K_i$ em que as funções $g_i(X)$ são lineares e as constantes K_i são reais. Os modelos de programação linear são um sub-conjunto do universo dos modelos de optimização. Quando as variáveis são todas inteiras, o modelo designa-se por modelo de Programação Inteira e quando algumas variáveis são inteiras e outras mistas, o modelo designa-se por modelo de Programação Inteira Mista. O estudo de algoritmos e/ou heurísticas para a resolução de problemas de

optimização faz parte de um domínio da Matemática designado por Investigação Operacional.

Esta tese endereça o dimensionamento de redes ATM para o suporte de múltiplos serviços em que o conjunto de serviços pode ser uma mistura de três tipos diferentes: serviços conversacionais, serviços de consulta e serviços ponto-multiponto. Os serviços conversacionais caracterizam-se por permitirem o estabelecimento de ligações ponto-a-ponto bidireccionais simétricas entre quaisquer dois clientes do serviço. Ligações bidireccionais simétricas são ligação que requisitam a mesma capacidade de comunicação em ambos os sentidos. Nos serviços de consulta (tradução do termo anglo-saxónico *retrieval services*), existem dois tipos diferentes de actores, os clientes e os servidores. Neste tipo de serviços, uma ligação ponto-a-ponto pode ser solicitada à rede de um cliente para um servidor e as ligações são habitualmente bi-direccionais assimétricas (com maior capacidade no sentido do servidor para o cliente). Finalmente, um serviço ponto-multiponto caracteriza-se por permitir a comunicação entre vários clientes de uma forma simultânea em que cada cliente estabelece uma ligação ponto-multiponto uni-direccional com os restantes clientes.

O dimensionamento de uma rede ATM deverá ter em conta os mecanismos de gestão de recursos que são usados na rede durante a sua operação. Os problemas estudados consideram uma gestão de recursos baseado em encaminhamento fixo multi-horário. Numa rede com encaminhamento fixo, existem trajectos únicos definidos à priori (ou árvores para ligações ponto-multiponto) entre os vários nós da rede pelos quais a rede estabelece os VCCs desde que haja os recursos necessários disponíveis. No caso do encaminhamento fixo constante, os trajectos definidos permanecem constantes ao longo de todo o tempo. No caso do encaminhamento fixo multi-horário, permite-se que os trajectos sejam alterados ao longo do tempo. Neste último caso, o tempo é particionado em intervalos de tempo pré-definidos e o trajecto de encaminhamento pode variar nos instantes entre intervalos de tempo. Nesta tese, por oposição à designação “multi-horário”, o encaminhamento constante será referido por encaminhamento uni-horário tendo subjacente a ideia que o encaminhamento é planeado para um único período horário.

A tecnologia ATM permite a configuração da rede com caminhos virtuais designados na terminologia da norma por VPCs (*Virtual Path Connections*). Um caminho virtual, ou VPC, permite encaminhar múltiplos VCCs de uma forma agregada. Um VCC ponto-a-

ponto pode ser encaminhado por um único VPC (se ele existir entre a origem e o destino da ligação) ou pode ser encaminhado por uma concatenação de diferentes VPCs. Um VPC extremo-a-extremo é um VPC que suporta apenas ligações (VCCs) entre os seus nós extremo. A configuração de redes lógicas exclusivamente constituídas por VPCs extremo-a-extremo apresenta algumas vantagens, das quais uma das mais importantes é a simplificação dos mecanismos de CAC e, consequentemente, a minimização dos tempos de estabelecimento de VCCs. A primeira parte do trabalho apresentado nesta tese considera esta solução para a gestão de recursos da rede e endereça o dimensionamento de redes ATM para o suporte de todos os tipos de serviços referidos anteriormente [Sousa99].

No entanto, uma rede lógica baseada apenas em VPCs extremo-a-extremo apresenta alguns inconvenientes. Primeiro, esta arquitectura não é facilmente escalável para redes físicas de maior dimensão devido ao número de VPCs que é necessário configurar na rede. Segundo, esta arquitectura não explora ganhos de partilha de recursos dado que cada VPC suporta apenas o fluxo de tráfego entre os seus nós extremo. Uma arquitectura para as redes lógicas que permita que o tráfego entre os nós da rede possa percorrer mais do que um VPC conduz a que diferentes fluxos de tráfego possam partilhar VPCs comuns e, consequentemente, haja ganhos de partilha de recursos. Assim, nesta tese é proposta uma nova arquitectura para as redes lógicas de suporte aos serviços conversacionais e são estudados modelos de dimensionamento adequados [Sousa01]. Esta nova arquitectura baseia-se em redes lógicas hierarquizadas a 2 níveis. A arquitectura proposta, para além de permitir explorar ganhos de partilha de recursos, reduz o número total de VPCs que é necessário configurar na rede e limita o número de VPCs que cada ligação atravessa a um máximo de três (um factor determinante para serviços comutados em que o tempo de estabelecimento de VCCs é proporcional a este valor).

Finalmente, uma arquitectura com VPCs extremo-a-extremo torna-se ainda menos eficaz para o caso dos serviços ponto-multiponto porque, para além das desvantagens descritas anteriormente, esta arquitectura também não explora a capacidade que os comutadores ATM têm de replicação das células ATM de uma entrada para várias saídas. Como alternativa, nesta tese é investigada a utilização de uma arquitectura em árvore para as redes lógicas de suporte a serviços ponto-multiponto e são estudados os modelos de dimensionamento adequados.

1.3 Evolução do Estado-da-Arte

Grande parte do trabalho publicado no passado no domínio da gestão de recursos e do dimensionamento de redes ATM baseia-se no desenho de redes lógicas constituídas por VPCs extremo-a-extremo. Esta estratégia tem sido proposta como solução para diferentes aspectos de gestão de recursos em redes ATM [Frisen96].

Em alguns casos é explorada a capacidade de reconfiguração das redes ATM e são propostas redes lógicas reconfiguráveis entre diferentes períodos de tempo de modo a melhor adaptarem os recursos alocados às variações de tráfego submetido à rede. Em [Medhi94], são propostos modelos de optimização lineares para o suporte a serviços conversacionais simétricos em que o autor considera o dimensionamento multi-horário da rede. O mesmo autor em [Medhi95] propõe a resolução dos modelos de dimensionamento da rede através de uma heurística baseada em relaxação Lagrangeana com optimização por sub-gradiente. Este trabalho contempla, para cada VPC extremo-a-extremo o cálculo prévio de caminhos candidatos através de um algoritmo de *k*-trajectos mais curtos.

O conceito de encaminhamento dinâmico foi explorado em primeiro lugar nas redes de comutação de circuitos [Comm95], quer como forma de tornar a rede mais robusta a períodos de pico de tráfego, quer como forma de optimizar a utilização dos recursos da rede, quer ainda como forma de dotar a rede de algum grau de recuperação a falhas. [Gupta92] [Hwang94] [Hwang95] comparam o desempenho de diferentes formas de encaminhamento dinâmico em redes lógicas com VPCs extremo-a-extremo. Em [Medhi97], é explorado o uso de encaminhamento dinâmico no dimensionamento de redes ATM. Neste caso, o problema é separado em 2 passos. No primeiro passo, são usadas aproximações com base no pressuposto que as probabilidades de bloqueio são pequenas (aproximadamente 1%) para estimar as capacidade de cada VPC assumindo o encaminhamento dinâmico. Tendo como base uma rede de 10 nós, é demonstrado por simulação que as aproximações são realistas para diferentes esquemas de encaminhamento dinâmico. No segundo passo, é usado o modelo anteriormente apresentado [Medhi95] para dimensionar a rede física.

Noutros casos, é considerado o dimensionamento de recursos suficientes de modo a que em caso de falha de um elemento da rede, o tráfego afectado possa ser reencaminhado na sua totalidade ou numa percentagem cujo valor é um parâmetro do problema [Girard98] [Iraschko98] [Murakami94] [Xiong99] [Al-Rumaih96] [Gersht94]. Em [Veitch97] é feito

um levantamento do estado-da-arte deste tipo de técnicas propostas na literatura para redes ATM.

Em [Kim95], é abordado o dimensionamento de redes lógicas com VPCs extremo-a-extremo em que a rede física é fixa. O problema abordado considera uma função objectivo dependente dos custos de capacidade, dos custos de ocupação das filas de espera e dos custos associados aos mecanismos de controlo da rede. O modelo de dimensionamento resultante é não-linear e é resolvido para uma rede de 6 nós com um método de programação não-linear. O mesmo autor em [Kim96] endereça as várias fases de planeamento de uma rede ATM: determinação das localizações dos nós da rede, determinação das capacidades de transmissão entre nós e determinação das ligações físicas entre nós. Para cada uma das fases, são propostos modelos de programação e são propostas heurísticas para a sua resolução.

Algumas referências mais antigas estudam o problema de dimensionamento de redes de computadores. Os primeiros trabalhos endereçam o desenho da topologia de redes em que o custo das ligações entre dois nós da rede é dado por um valor fixo se as ligações existirem na solução de dimensionamento (referido na literatura como *uncapacitated networks*). Este problema é mais fácil de resolver porque o intervalo dual entre a solução óptima do problema de dimensionamento e a sua relaxação linear é pequeno [Balakrishnan89]. No caso em que o dimensionamento tem como objectivo saber quantas ligações físicas e de que tipo se devem instalar entre os nós da rede (referido na literatura como *capacitated networks*), o intervalo dual é muito maior e as técnicas tradicionais não se aplicam pela sua ineficiência [Bienstock96], [Magnanti95], [Gavish89] [Gavish90] [LeBlanc89] [Ng87] [Maruyama77]. Por exemplo, [Magnanti95] endereça o problema de dimensionamento de redes em que é possível instalar 2 tipos diferentes de ligações entre quaisquer nós da rede. São propostos 3 tipos diferentes de restrições que melhoram o valor da relaxação linear do problema entre 55% e 80% para redes entre 6 nós e 15 nós e obtém um intervalo dual médio de 8.13%. Ainda assim, este método só consegue resolver problemas até à optimalidade para redes até 15 nós. Em [Bienstock96], é estudado o mesmo problema com uma formulação diferente e são também propostos 3 outros tipos de restrições que, quando adicionadas ao problema original, melhoram o valor da relaxação linear do problema. Os resultados são melhores que os de [Magnanti95] em termos de

tempo de cálculo da solução óptima mas o autor foi apenas capaz de resolver problemas para redes até 16 nós e 49 arestas.

Alguns autores propõem diferentes metodologias no âmbito do desenho de redes lógicas com VPCs não extremo-a-extremo, embora o número de referências neste domínio seja bem menor que no caso anterior. Toda a literatura referenciada considera apenas tráfego ponto-a-ponto simétrico e em nenhum caso é endereçado o problema de dimensionamento da rede com o objectivo de minimização do seu custo (o problema abordado nesta tese). Em alguns casos, o objectivo é minimizar alguns parâmetros considerados como boas características no dimensionamento de uma rede (por exemplo, minimização da carga máxima de cada ligação). Noutros, a rede é fixa e o desenho das redes lógicas é feito com o objectivo de maximizar parâmetros de desempenho ou de rendimento económico da rede. Uma metodologia comum a quase todos os trabalhos referenciados é, face à complexidade dos modelos de optimização resultantes, a decomposição da sua solução em diferentes passos de modo a separar o problema global em sub-problemas mais fáceis de resolver, penalizando a qualidade das soluções obtidas. De seguida, são descritos alguns exemplos ilustrativos.

Em [Chlamtac93] e [Chlamtac94], é apresentada uma metodologia que determina os VPCs a configurar na rede e os fluxos a atribuir a cada VPC de modo a que a carga máxima de cada ligação seja minimizada. Um fluxo entre 2 nós da rede é dado por um valor de capacidade. A metodologia proposta comporta 3 fases: (1) determinação dos nós extremo de cada VPC, (2) atribuição dos fluxos a cada VPC e (3) determinação do percurso de cada VPC. [Ahn94] endereça um problema semelhante permitindo que o tráfego entre dois nós de rede possa ser encaminhado por diferentes percursos (encaminhamento alternativo). Usa também uma metodologia em três passos: (1) determinação das capacidades para cada um dos percursos alternativos, (2) determinação de caminhos alternativos que respeitem as restrições de atraso e de capacidade das ligações físicas e (3) determinação do conjunto de VPCs capaz de suportar as capacidades por todos os percursos alternativos calculados anteriormente. Em [Lin93] e [Cheng94], o objectivo é minimizar a probabilidade de bloqueio da rede sujeito a restrições de tempo máximo de estabelecimento de chamada (dado pelo número máximo de VPCs atravessados). Este caso contempla o tráfego de serviços comutados e é formulado através de um modelo de optimização combinatória não linear e não diferenciável. É proposta uma solução

heurística em duas fases. Na primeira fase, a capacidade de cada VPC é fixada e a atribuição de fluxos é ajustada. Na segunda fase, as capacidades de cada VPC são ajustadas usando a atribuição de fluxos dada pela fase anterior.

Em [Sethi95], é proposta uma arquitectura de múltiplas árvores abrangentes (uma para cada nó origem) para o suporte a serviços conversacionais. As capacidades dos VPCs de cada árvore são calculadas previamente para cada possível árvore abrangente com base no tráfego e, em seguida, é formulado um problema de optimização linear para o desenho das várias árvores abrangentes. Neste problema, as capacidades das ligações físicas é fixa e os resultados, calculados para redes até 7 nós, mostram que é possível para a mesma rede conseguir melhores valores de probabilidade de bloqueio do que com VPCs extremo-a-extremo. Esta abordagem tem no entanto alguns inconvenientes. Primeiro, usa a noção de Árvore Virtual (*Virtual Tree*) que não é suportado pela norma ATM. Segundo, permite que as árvores abrangentes de diferentes nós origem passem por diferentes ligações físicas da rede. Este facto leva a que o estabelecimento de uma chamada ponto-a-ponto bidireccional seja feito por diferentes trajectos nos dois sentidos resultando em alguns problemas operacionais (maior tempo de estabelecimento de ligação, maior impacto na convergência de um processo de recuperação a falhas, etc…).

Em [Cidon94] [Gerstel96], é proposto um algoritmo recursivo que determina a configuração das redes lógicas em que é possível atingir um compromisso entre o tempo de estabelecimento de chamadas (dado pelo número máximo de VPCs atravessado por cada par origem-destino) e a utilização das tabelas de encaminhamento dos comutadores (dado pelo número de VPCs que partilham cada ligação física da rede). O algoritmo é concebido para redes lógicas em árvore de suporte mas não leva em linha de conta o custo da rede física resultante como parâmetro de optimização.

Em [Mitra96] [Gibbens95] [Anerousis98] [Faragó95] são propostas diferentes alternativas para a determinação de redes lógicas com VPCs não extremo-a-extremo em que a configuração física da rede é dada e o objectivo é maximizar o rendimento económico da rede. As classes de tráfego são também caracterizadas pelo rendimento económico que o operador obtém em as suportar. Os problemas são formulados por modelos de optimização não linear e não ambicionam a determinação de nenhuma arquitectura especial para as redes lógicas. Embora não sejam problemas de dimensionamento da rede, estes exemplos ilustram também o grau de complexidade

resultante da determinação conjunta da capacidade e do percurso dos VPCs numa arquitectura com VPCs não extremo-a-extremo.

Alguns investigadores argumentam que a separação dos diferentes serviços em diferentes redes lógicas diminui os ganhos de multiplexagem estatística que a tecnologia ATM permite obter. Por outro lado, a mistura de diferentes serviços na mesma rede lógica exige que a Qualidade de Serviço suportada pela rede seja dimensionada para o serviço mais exigente. Assim, algumas referências exploram este facto para o dimensionamento da rede. [Sato91] propõe um metodo de dimensionamento que leva em linha de conta a caracterização estatística das fontes de tráfego de forma a determinar aproximadamente as capacidades das classes de trafego resultantes. [Gerla89] estuda o desenho da rede de transporte de suporte a tráfego ATM em que os comutadores ATM definem os pares origem-destino das capacidades a instalar na rede. [Evans91] dimensiona um conjunto de VPCs a serem partilhados por todos os serviços com o objectivo de maximizar o desempenho da rede.

Finalmente, de notar que não foi encontrada nenhuma referência na literatura que enderece o problema do dimensionamento de redes para suporte de serviços ponto-multiponto. O trabalho publicado referente ao suporte de serviços ponto-multiponto insere-se no tema da gestão de recursos em que o problema endereçado é como determinar o encaminhamento de ligações ponto-multiponto em redes de telecomunicações. O capítulo 6 de [Dziong97] e todas as referências nele contidas são uma excelente referência do trabalho publicado neste domínio para redes ATM.

1.4 Estrutura da Tese

O segundo capítulo endereça os aspectos que servem de base à formulação dos diferentes problemas de dimensionamento abordados nos capítulos seguintes e que são comuns a mais do que um deles. Começa por apresentar o modelo da rede física e do seu custo como medida de qualidade adoptada em todos os problemas abordados. Apresenta também os modelos de tráfego adoptados tanto ao nível da ligação lógica como ao nível da célula e justifica os métodos usados para a geração aleatória de cenários de tráfego. Neste capítulo são também apresentadas duas alternativas para os modelos das redes lógicas: modelização explícita de VPCs e modelização por trajectos candidatos. É ainda abordada a definição de probabilidade de bloqueio em serviços comutados, definição esta que tem

impacto na forma como são determinadas as capacidades dos VPCs que formam as redes lógicas de suporte. Finalmente, é abordado o aspecto da integralidade dos VPCs que habitualmente não é considerado nos modelos de dimensionamento de redes ATM; é proposta uma forma de incluir as restrições de integralidade dos VPCs num exemplo de um problema de dimensionamento e, com base na sua análise, é justificada a razão porque nos modelos de dimensionamento dos capítulos seguintes este aspecto não é considerado.

O terceiro capítulo endereça o dimensionamento de redes ATM para o suporte simultâneo dos diferentes tipos de serviços assumindo a configuração de redes lógicas de suporte baseadas exclusivamente em VPCs extremo-a-extremo. Esta topologia permite separar o cálculo das capacidades dos VPCs da determinação dos seus trajectos. Assim, o capítulo começa por abordar a forma como as capacidades dos VPCs são calculadas para os diferentes tipos de serviços abordados. De seguida, são apresentados dois modelos de Programação Inteira que definem o problema de dimensionamento com base nas duas alternativas para os modelos das redes lógicas introduzidas no capítulo anterior: modelização explícita de VPCs e modelização por trajectos candidatos. De seguida, é proposto um algoritmo heurístico baseado na técnica da Relaxação Lagrangeana com Optimização por Sub-Gradiente que permite resolver os dois modelos apresentados. Finalmente, são apresentados resultados computacionais e é feita uma avaliação da eficiência e da eficácia do algoritmo proposto. Com base nos resultados computacionais, é feita também uma avaliação das duas alternativas de modelização das redes lógicas e é ilustrada a utilidade da utilização de modelos multi-horário com base nos ganhos que se obtêm comparativamente com os modelos uni-horário.

No quarto capítulo, é abordado o problema de dimensionamento de redes ATM configuradas com redes lógicas hierarquizadas a dois níveis para o suporte de serviços conversacionais. O capítulo começa por apresentar a arquitectura proposta e descrever as suas vantagens operacionais relativamente às redes lógicas com VPCs extremo-a-extremo. De seguida, são descritas e justificadas as aproximações necessárias que tornam possível combinar de uma forma linear o cálculo das capacidades dos VPCs com o dimensionamento da rede física. São então apresentados modelos de Programação Inteira que definem o problema de dimensionamento. Finalmente, são apresentados resultados computacionais em que os modelos propostos são resolvidos pelo algoritmo de *branch-and-bound*. Pelos resultados experimentais apresentados, é feito um estudo comparativo

entre o dimensionamento das redes ATM para o suporte de serviços conversacionais com VPCs extremo-a-extremo e com a nova arquitectura proposta.

O quinto capítulo aborda o dimensionamento de redes ATM para serviços ponto-multiponto suportados por redes lógicas em árvore. Neste caso não é possível separar o cálculo das capacidades dos VPCs da determinação dos seus trajectos. Assim, o capítulo começa por apresentar um método que permite modelizar o cálculo das capacidades dos VPCs através de restrições lineares em que se garante que a probabilidade de bloqueio pretendida é cumprida. Pelo método proposto, são apresentadas duas formas diferentes para as restrições lineares resultantes denominadas por modelo binário e modelo real. De seguida, são apresentados dois modelos de dimensionamento baseados nas duas formas de modelização das capacidades dos VPCs. Com base nestes dois modelos, denominados modelos naturais para o problema endereçado, o capítulo apresenta também duas reformulações válidas cujo objectivo é tornar mais eficiente a sua resolução. Finalmente são apresentados resultados computacionais que permitem avaliar a eficiência dos modelo naturais e reformulados. Com base nos resultados apresentados, é feito também um estudo comparativo entre o dimensionamento das redes ATM para o suporte de serviços ponto-multiponto com VPCs extremo-a-extremo e com a nova arquitectura proposta.

Finalmente, as principais conclusões do trabalho descrito nesta tese são apresentadas no sexto capítulo. Nele são também apontados os tópicos deixados em aberto pelo presente trabalho e considerados mais importantes para trabalho futuro.

1.5 Principais Contribuições

- Proposta de dois modelos de Programação Inteira para o dimensionamento de redes ATM com VPCs extremo-a-extremo no suporte simultâneo a serviços conversacionais, serviços de consulta e serviços ponto-multiponto. Os modelos propostos extendem modelos previamente conhecidos de modo a poder ser considerado tráfego assimétrico (característico dos serviços de consulta e serviços ponto-multiponto) e modelização de VPCs em que um dos extremos não é conhecido à partida (característica dos serviços de consulta quando existem múltiplos servidores na rede).
- Proposta de um algoritmo heurístico baseado na técnica de Relaxação Lagrangeana com Optimização por Sub-Gradiente para a resolução dos modelos propostos para

redes com VPCs extremo-a-extremo. Estudo da influência da escolha entre modelização explícita de VPCs e modelização por trajectos candidatos no desempenho do algoritmo. Estudo dos principais factores que influenciam a qualidade das soluções encontradas pelo algoritmo.

- Proposta de uma arquitectura hierárquica com dois níveis para as redes lógicas de suporte a serviços conversacionais. Determinação de modelos de dimensionamento adequados à arquitectura proposta. Estudo comparativo entre o dimensionamento das redes ATM com VPCs extremo-a-extremo e com a nova arquitectura proposta.
- Proposta de uma arquitectura em árvore para as redes lógicas de suporte a serviços ponto-multiponto. Determinação de modelos de dimensionamento adequados à arquitectura proposta. Estudo comparativo entre o dimensionamento das redes ATM com VPCs extremo-a-extremo e com a nova arquitectura proposta.
- Proposta de um método de geração aleatória de tráfego ponto-multiponto que permite escolher perfis de tráfego com diferentes valores de número médio e número máximo de destinos. Este método permite estudar a influência de diferentes perfis de tráfego no estudo comparativo entre o dimensionamento das redes ATM com VPCs extremo-a-extremo e com redes lógicas em árvore.
- Proposta de formas lineares de formulação de restrições de probabilidade de bloqueio, ao nível da ligação lógica, na determinação das capacidades dos VPCs. Este método permite derivar modelos lineares de dimensionamento da rede quando não é possível separar o cálculo das capacidades dos VPCs da determinação dos seus trajectos.

CAPÍTULO 2

CONCEITOS BASE DE DIMENSIONAMENTO

Para que seja possível formular um problema de dimensionamento, é necessário primeiro endereçar separadamente a modelização dos diferentes aspectos que estão em jogo no dimensionamento de uma rede ATM. Este capítulo endereça os aspectos que servem de base à formulação dos diferentes problemas de dimensionamento abordados nos capítulos seguintes e que são comuns a mais do que um deles. O capítulo começa por apresentar o modelo da rede física e do seu custo como medida de qualidade adoptada em todos os problemas de dimensionamento abordados. Depois são apresentadas formas alternativas para a modelização dos trajectos dos VPCs que compõem cada uma das redes lógicas. São também apresentados os modelos de tráfego adoptados tanto ao nível da ligação lógica como ao nível da célula bem como os métodos usados para a geração aleatória de cenários de tráfego. É ainda abordada a definição de probabilidade de bloqueio em serviços comutados, definição esta que tem impacto na forma como são determinadas, nos capítulos seguintes, as capacidades dos VPCs que formam as redes lógicas de suporte. Finalmente, é discutido o aspecto da integralidade dos VPCs que habitualmente não é considerado nos modelos de dimensionamento de redes ATM. Para isso, é proposta uma forma de incluir as restrições de integralidade dos VPCs num exemplo de um problema de dimensionamento de uma rede ATM.

2.1 Modelo da Rede Física

Esta tese considera como parâmetros de entrada do problema do dimensionamento das redes ATM, o conjunto dos possíveis nós de comutação da rede e, para cada par de nós, os tipos de ligação e número máximo de ligações de cada tipo que é possível pôr em operação. Por exemplo, uma especificação típica do modelo da rede física poderá ser: a rede tem 3 nós de comutação possíveis (A, B e C); entre os nós A e B podem ser usadas no máximo 4 ligações SDH STM-1 (155 Mbps) e 2 ligações SDH STM-4 (622 Mbps); entre os nós A e C podem ser usadas no máximo 8 ligações PDH E3 (34 Mbps). Um dos resultados do dimensionamento da rede é a determinação dos nós que devem pertencer à rede e das ligações (tipo e número) que devem existir entre os nós seleccionados.

Considere-se N o conjunto de todos os possíveis nós de comutação da rede a dimensionar. Considere-se também A o conjunto de todos os pares de nós entre os quais é possível implementar pelo menos um dos tipos de ligações físicas. Estes dois conjuntos definem um grafo não orientado (N,A) em que N é o conjunto dos nós e A é o conjunto das arestas do grafo. Ao longo de toda a tese, as letras i e j são usadas para representar os elementos de N e a dimensão de N é dada por $|N|$. Os elementos de A são representados por pares não ordenados na forma $\{i,j\}$ e a dimensão de A é dada por $|A|$. Sempre que é necessário especificar o sentido de cada elemento de A, (i,j) representa a aresta $\{i,j\}$ de i para j e (j,i) representa a aresta $\{i,j\}$ de j para i. No exemplo da Figura 2-1, $N = \{1,2,3,4,5\}$, $A = \{\{1,2\},\{1,3\}, \{1,5\},\{2,3\},\{2,4\},\{4,5\}\}$, $|N| = 5$ e $|A| = 6$.

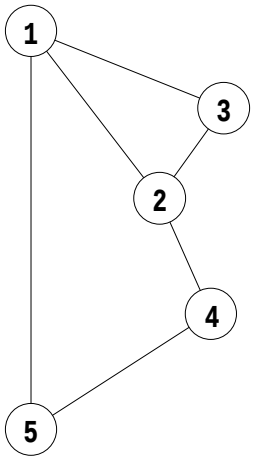


Figura 2-1 : Exemplo de um grafo representativo de uma rede com 5 nós e 6 arestas

Considera-se também o conjunto T de todos os tipos de ligações físicas que é possível pôr em operação na rede. Um elemento de T é representado por t e a sua capacidade de transmissão (também designada por largura de banda) é representada por α_t. Por exemplo, considere-se que para a rede da Figura 2-1 é possível usar os seguintes tipos de ligação: PDH E3 (34 Mbps), SDH STM-1 (155 Mbps) e SDH STM-4 (622 Mbps). Uma forma possível de modelizar esta informação é:

$T = \{1,2,3\}$ $\quad t = 1 \Rightarrow$ PDH E3 $\quad \alpha_1 = 34E+6$

$\quad t = 2 \Rightarrow$ SDH STM-1 $\quad \alpha_2 = 155E+6$

$\quad t = 3 \Rightarrow$ SDH STM-4 $\quad \alpha_3 = 622E+6$

Define-se $Y^t_{\{ij\}}$ como sendo o número máximo de ligações do tipo t que é possível pôr em operação na aresta $\{i,j\}$. A rede física resultante do processo de dimensionamento é modelizada pelas variáveis inteiras não negativas $y^t_{\{ij\}}$ que representam o número de ligações físicas do tipo t na aresta $\{i,j\}$. Nesta tese, designaremos por variáveis y o conjunto de todas as variáveis deste tipo de um modelo de dimensionamento. A Figura 2-2 ilustra a rede física resultante para uma solução em que as variáveis y são dadas por:

$y^1_{\{12\}} = 2$, $y^1_{\{23\}} = y^2_{\{24\}} = y^1_{\{45\}} = 1$

$y^t_{\{ij\}} = 0$ para todos os outros valores de t, i e j

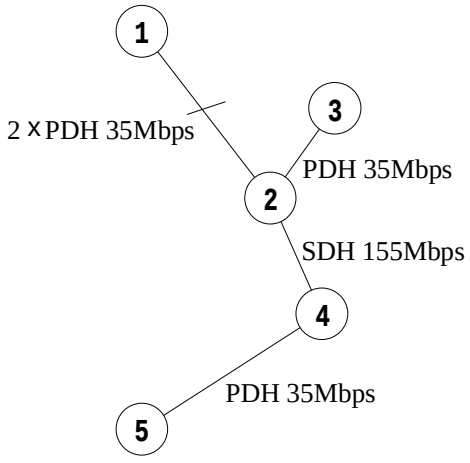


Figura 2-2 : Exemplo de uma solução para a rede física

Segundo a notação introduzida, a capacidade de transmissão total da aresta $\{i,j\}$ é uma função linear das variáveis y dada por:

$$\sum_{t \in T} \alpha_t y^t_{\{ij\}} \qquad (1)$$

Note-se que não é necessário criar nenhum tipo de variável que identifique explicitamente os elementos de *N* que pertencem à rede resultante. Esta informação está implícita no valor das variáveis *y*. O elemento *i* de *N* pertence à solução se:

$$\sum_{j \in N : \{i,j\} \in A} \sum_{t \in T} y^t_{\{ij\}} > 0 \qquad (2)$$

2.2 Modelo do Custo da Rede Física

O objectivo do trabalho apresentado nesta tese é a determinação da rede física que minimiza o seu custo de operação e manutenção. É necessário modelizar o custo de uma solução de forma apropriada de modo a reflectir o custo real da rede física. Podemos distinguir dois tipos de custos referentes aos dois tipos de elementos que uma rede comporta: os custos associados à transmissão da informação pelas ligações físicas e os custos associados ao encaminhamento da informação em cada nó de comutação. As secções 2.2.1 e 2.2.2 abordam os dois tipos de custos separadamente e na secção 2.2.3, é apresentado o modelo de custos completo adoptado nesta tese.

2.2.1 Custos de Transmissão

Os custos de transmissão são os custos que o operador da rede tem de suportar para transmitir células ATM entre quaisquer dois comutadores da rede. Estes custos dependem da forma como o operador implementa estas ligações. Podem-se distinguir 3 casos diferentes: ligações físicas dedicadas, ligações de rede ou ligações contratadas.

No caso das ligações físicas dedicadas, o operador tem no terreno cabos (por exemplo, de fibra óptica) que interligam os locais dos possíveis nós de comutação que podem ser usados para implementar as ligações entre os comutadores. Neste caso, o operador deverá saber avaliar qual o custo de operação e manutenção de uma infraestrutura física de comunicações e, deste valor, deverá ser possível determinar um custo de operação e manutenção C_l por unidade de comprimento (por exemplo, por kilómetro) de cada par de fibras ópticas. Para se saber o custo de cada ligação, é necessário que o operador especifique o comprimento (por exemplo, em kilómetros) das fibras que ligam cada par de nós, designado por $l\{i,j\}$.

No caso das ligações físicas serem suportadas por uma rede de transmissão (por exemplo, uma ligação SDH 155 Mbps pode ser suportada por uma rede de tecnologia SDH), o operador deverá saber avaliar quanto lhe custa suportar cada tipo de ligação na rede de transmissão respectiva. Designamos por $C_t\{i,j\}$ o custo de suportar uma ligação física do tipo *t* na rede de transmissão entre os possíveis nós de comutação *i* e *j*.

Finalmente, no caso de ligações contratadas, este custo representa o preço cobrado por cada ligação pelo terceiro operador. Designamos por $C_t^S\{i, j\}$, o preço a pagar por uma ligação dedicada do tipo *t* entre os locais dos possíveis nós de comutação *i* e *j*.

2.2.2 Custos de Comutação

Os custos de comutação são os custos que o operador tem de suportar para manter em funcionamento os nós de comutação. Um modelo fiel destes custos é bastante complexo. Os comutadores ATM comercialmente disponíveis são de dois tipos: comutadores de configuração fixa e comutadores escaláveis. Os primeiros são tipicamente comutadores de pequena capacidade mais endereçados ao mercado das redes de área local (LANs – *Local Area Networks*) e que os operadores públicos habitualmente usam como comutadores de acesso. Os comutadores escaláveis têm uma configuração modular que permite ao operador adquirir apenas o número de módulos suficiente para suportar o tráfego que necessita. Desta forma, é possível rentabilizar melhor os investimentos sem que haja necessidade de investimentos iniciais tão elevados.

Um modelo completo dos custos de comutação deveria contemplar um valor fixo para cada nó (inerente a toda a infraestrutura necessária para ter pelo menos um módulo em operação tal como espaço físico, energia eléctrica, etc…) mais um valor proporcional ao número de módulos instalados no nó de comutação (inerente a todo o processo de operação e manutenção dos equipamentos de comutação). No entanto, um modelo de custos com esta estrutura aumenta bastante a complexidade do problema de dimensionamento porque exige que se modelize o número de módulos na formulação resultante e este é dependente do número e tipo de ligações físicas necessárias em cada nó e varia de fabricante para fabricante.

Para obviar esta dificuldade, é habitual dimensionar uma rede ATM separando o dimensionamento das ligações físicas do dimensionamento dos nós de comutação. Esta é a aproximação considerada nesta tese. Os problemas de dimensionamento propostos

endereçam o dimensionamento das ligações físicas e pressupõem um pós-processamento dos equipamentos de comutação necessários instalar em cada nó de comutação. No entanto, ignorar completamente os custos de comutação pode não ser aceitável nos casos em que eles não são insignificantes quando comparados com os custos de transmissão. O modelo adoptado nesta tese assume que o operador atribui um custo de comutação C_t^i por cada tipo de interface *t* em cada nó de comutação *i*. O operador deverá saber avaliar qual o o custo de operação e manutenção do nó de comutação *i* capaz de encaminhar uma capacidade de informação de 34 Mbps, 155 Mbps, 622 Mbps, etc.., tantos valores quantos os tipos de ligações físicas que pretende usar. Assim, o custo de comutação de nó *i* é igual à soma dos custos associados à capacidade de cada ligação física desse nó.

2.2.3 Modelo de Custos Adoptado

Designa-se por $C_{\{ij\}}^t$ o custo de operação e manutenção associado à utilização de uma ligação do tipo *t* entre os nós de comutação *i* e *j*. Estes valores são dados por:

$$C_{\{ij\}}^t = C_t^i + C_t^j + \begin{cases} C_l \cdot l\{i,j\} & \text{, se ligação física dedicada} \\ C_t\{i,j\} & \text{, se ligação de rede} \\ C_t^s\{i,j\} & \text{, se ligação contratada} \end{cases} \quad (3)$$

Os dois primeiros termos referem-se aos custos de comutação nos dois nós extremos da ligação e o último termo refere-se aos custos de transmissão que são dependentes da forma como as ligações são implementadas. Segundo a notação introduzida, o custo total de uma rede física ATM é uma função linear das variáveis *y* dada por:

$$\sum_{\{i,j\} \in A} \sum_{t \in T} C_{\{ij\}}^t \cdot y_{\{ij\}}^t \quad (4)$$

Esta expressão é a função objectivo que se pretende minimizar em todos os problemas de dimensionamento abordados nesta tese.

2.3 Modelo de uma Rede Lógica

Esta tese considera o dimensionamento de redes ATM multi-serviço em que cada serviço é suportado por uma rede lógica. Cada rede lógica é constituída por um conjunto de VPCs. Assim, a modelização de uma rede lógica é equivalente à modelização de cada um dos VPCs que a compõem. Para cada serviço, designam-se por nós de tráfego os nós da

rede aos quais o serviço é oferecido e designam-se por nós de trânsito os nós que não são de tráfego. Os VPCs de uma rede lógica são configurados na rede física interligando os nós de tráfego do serviço suportado pela rede lógica.

Considere S o conjunto das redes lógicas s a configurar na rede. Considere K_s o conjunto dos VPCs k que compõem a rede lógica s. Para cada VPC é necessário calcular a capacidade em cada um dos seus dois sentidos e o trajecto na rede física entre os seus dois nós extremo. Atribuíndo uma origem e um destino aos extremos de cada VPC, o valor b_k^s designa a capacidade do VPC k pertencente à rede lógica s no sentido da origem para o destino e $\underline{b}_k^s$, no sentido do destino para a origem. A necessidade de definir de entre os nós extremos de um VPC qual a origem e qual o destino prende-se com o facto de no caso geral um VPC poder ter diferentes capacidades nos seus dois sentidos. As capacidades dos VPCs que constituem uma rede lógica dependem: (i) da topologia da rede lógica, (ii) do tráfego ao nível da ligação lógica do serviço suportado e (iii) do tráfego ao nível da célula do serviço suportado. Nesta secção é abordado o problema de como modelizar o trajecto de cada VPC. Consideramos dois tipos de aproximação: modelização explícita e modelização por trajectos candidatos. Cada uma tem as vantagens e desvantagens que passamos a descrever de seguida.

2.3.1 Modelização Explícita de VPCs

O caso mais simples é quando os nós extremos do VPC são conhecidos à partida; classificamos estes como VPCs com extremos definidos. Um VPC $k \in K_s$ deste tipo é caracterizado pelo seu nó origem o_k^s, pelo seu nó destino d_k^s e por um trajecto modelizado pelas variáveis binárias x_{ij}^{ks}. Quando esta variável é 1, ela define que o trajecto da origem para o destino do VPC $k \in K_s$ passa em (i,j), ou seja, na aresta $\{i,j\}$ no sentido de i para j. Nesta tese, designaremos por variáveis x o conjunto de todas as variáveis deste tipo. A designação de modelização explícita do VPC refere-se ao facto do trajecto ser modelizado pelo conjunto de variáveis binárias x cujo resultado o identifica explicitamente. Para isso, é necessário que as variáveis associadas ao VPC k tenham as seguintes relações entre si:

$$\sum_{j \in N:\{i,j\} \in A} \left(x_{ij}^{ks} - x_{ji}^{ks} \right) = 1 \quad \text{, para } i = o_k^s \tag{5a}$$

$$\sum_{j \in N:\{i,j\} \in A} \left(x_{ji}^{ks} - x_{ij}^{ks} \right) = 1 \quad \text{, para } i = d_k^s \tag{5b}$$

$$\sum_{j \in N:\{i,j\} \in A} \left(x_{ij}^{ks} - x_{ji}^{ks} \right) = 0 \quad \text{, para } i \in N \setminus \{ o_k^s , d_k^s \} \tag{5c}$$

A equação (5a) garante que existe pelo menos um arco de saída do nó origem. A equação (5b) garante que existe pelo menos um arco de entrada do nó destino. A equação (5c) garante que para todos os outros nós da rede, se houver um arco de entrada no trajecto, também há um arco de saída incluído no trajecto. Este conjunto de equações garante que o conjunto de variáveis cujo resultado é 1 inclui um trajecto do nó origem para o nó destino.

O segundo caso de interesse é quando um dos nós extremo (ou os dois) não é conhecido à partida mas pode ser um de entre um sub-conjunto de todos os nós da rede. Define-se região origem O_k^s (região destino D_k^s) como o conjunto dos nós pertencentes a *N* candidatos a serem a origem (o destino) do VPC *k* pertencente à rede lógica *s*. Neste caso, as variáveis *x* deverão respeitar as seguintes equações:

$$\sum_{i \in O_k^s} \sum_{j \in N:\{i,j\} \in A} \left(x_{ij}^{ks} - x_{ji}^{ks} \right) = 1 \tag{6a}$$

$$\sum_{i \in D_k^s} \sum_{j \in N:\{i,j\} \in A} \left(x_{ji}^{ks} - x_{ij}^{ks} \right) = 1 \tag{6b}$$

$$\sum_{j \in N:\{i,j\} \in A} \left(x_{ij}^{ks} - x_{ji}^{ks} \right) = 0 \quad \text{, para } i \in N \setminus \{ O_k^s , D_k^s \} \tag{6c}$$

A equação (6a) impõe que o número de arcos de saída em todos os nós de O_k^s seja mais um do que o número de arcos de entrada pelo que o trajecto do VPC deverá iniciar-se em qualquer um dos nós de O_k^s e poderá visitar outros nós de O_k^s sem que aí possa terminar. De forma análoga, a equação (6b) impõe que o trajecto do VPC deva terminar em qualquer um dos nós de D_k^s e poderá visitar os outros nós de D_k^s sem que aí possa começar. Finalmente, a equação (6c) garante a continuidade do trajecto nos nós não pertencentes a O_k^s nem a D_k^s. As equações (5) são um caso particular das equações (6) quando os conjuntos origem e destino são constituídos por um único nó cada.

2.3.2 Modelização de VPCs por Trajectos Candidatos

Neste tipo de modelização, para cada VPC é calculado à priori um conjunto de trajectos candidatos e a determinação do trajecto na solução é feita por escolha de um entre os trajectos deste conjunto. Considera-se P_k^s o conjunto dos trajectos candidatos *p* para o

VPC *k* pertencente à rede lógica *s*. Consideram-se também as constantes binárias δ_{ij}^{pks} que quando são 1 indicam que o trajecto candidato $p \in P_k^s$ do VPC *k* pertencente à rede lógica *s*, passa em (*i*,*j*). Cada VPC é modelizado pelas variáveis x_p^{ks} que definem se o trajecto candidato *p* é o trajecto solução do VPC *k* na rede lógica *s*. Segundo a notação apresentada, as variáveis *x* referentes a cada VPC *k* devem então obedecer à seguinte equação:

$$\sum_{p \in P_k^s} x_p^{ks} = 1 \tag{7}$$

Um aspecto importante neste tipo de modelização é a determinação do conjunto dos trajectos candidatos P_k^s para cada VPC (e que definem também os valores de δ_{ij}^{pks}). Se os conjuntos P_k^s não incluirem todos os trajectos possíveis entre os nós extremo do VPC *k*, então o modelo exclui à partida algumas das soluções possíveis do problema de dimensionamento. Neste caso, a utilização deste modelo no dimensionamento da rede não garante que a melhor solução calculada seja a solução óptima, porque não é possível saber à partida se a solução óptima é uma das soluções que foram excluídas como soluções possíveis.

O número de elementos de cada conjunto P_k^s é um parâmetro que permite obter um compromisso entre o tempo de cálculo da resolução do modelo de dimensionamento de rede e a qualidade da solução encontrada. Para instâncias do problema grandes, um modelo explícito dos VPCs (apresentado na secção anterior) poderá conduzir a tempos de cálculo demasiado longos e a adopção de um modelo por trajectos candidatos com uma dimensão pequena para os conjuntos P_k^s pode permitir determinar uma solução em tempo útil embora possivelmente sub-óptima.

Após a decisão de qual a dimensão $|P_k^s|$ a adoptar para o modelo, outro aspecto importante é que trajectos candidatos devem ser considerados e como podem ser determinados. Uma alternativa é o cálculo aleatório dos trajectos. No entanto, a experiência mostra que numa boa solução a maioria dos VPCs são encaminhados por um dos trajectos que atravessam o menor número possível de ligações físicas. Assim, é habitual determinar o conjunto P_k^s com os $|P_k^s|$ trajectos mais curtos em termos de número de saltos no grafo (*N*,*A*). Para VPCs com extremos definidos, existem algoritmos de cálculo de *k*-trajectos mais curtos que resolvem instâncias bastante grandes em tempo útil [Martins97]. No caso dos VPCs com extremos indefinidos, são calculados os $|P_k^s|$ trajectos

mais curtos entre todos os pares possíveis de nós origem-destino e, de entre estes, são escolhidos os $|P_k^s|$ trajectos com menor número de saltos.

Um modelo de VPCs por trajectos candidatos permite facilmente incluir restrições impostas pelo atraso máximo (parâmetro de Qualidade de Serviço) que as células ATM podem sofrer nos serviços suportados por uma determinada rede lógica. Este problema põe-se no dimensionamento de redes ATM de âmbito geográfico grande ou quando existem ligações físicas que usam ligações via satélite por exemplo. Nestes casos, o tempo de propagação de cada célula ATM nas ligações físicas pode ser significativo pelo que o trajecto a escolher para o estabelecimento de uma ligação lógica deve garantir o valor máximo de atraso permitido. Sabendo quais são os valores de atraso de todas as ligações físicas e os valores máximos de atraso em cada um dos nós de comutação, é fácil determinar quais dos trajectos, calculados por um algoritmo de cálculo de trajectos mais curtos, podem ser incluídos nos conjuntos P_k^s.

Um modelo explícito de VPCs que contemple restrições de atraso deve necessariamente incluir outras restrições para além das especificadas nas equações (5) e (6) o que torna o modelo resultante mais complexo e, consequentemente, mais difícil de resolver. Este aspecto não foi abordado no trabalho realizado e a sua referência é feita apenas porque os modelos apresentados em que os VPCs são modelizados por trajectos candidatos podem ser usados para o dimensionamento de redes com restrições de atraso (nos casos reais em que isso se justifique).

2.4 Modelo de Tráfego Oferecido

Uma rede ATM é simultaneamente uma rede orientada à ligação e uma rede de comutação de células. Assim, o cálculo das capacidades dos VPCs deve ser feito respeitando a qualidade de serviço contratada tanto ao nível da ligação como ao nível da célula. A capacidade dos VPCs depende também da topologia das redes lógicas pelo que o seu cálculo será abordado em cada um dos capítulos seguintes pois cada um deles propõe uma topologia diferente para as redes lógicas. No entanto, é necessário adoptar um modelo de tráfego que seja comum a todos os casos de dimensionamento para que estes possam ser comparados entre si. As secções seguintes endereçam separadamente a modelização do

tráfego ao nível da ligação (secção 2.4.1 para o tráfego ponto-a-ponto e secção 2.4.2 para o tráfego ponto-multiponto) e ao nível da célula (secção 2.4.3).

2.4.1 Nível da Ligação de Tráfego Ponto-a-Ponto

Os tipos de serviços abordados nesta tese contemplam dois casos de interesse. O primeiro caso é o dos serviços conversacionais em que o tráfego é bidireccional. O tráfego é caracterizado com base em pares origem-destino. Para cada serviço suportado pela rede, define-se o tráfego oferecido à rede de cada nó origem para cada nó destino. Um nó de tráfego de um serviço é um nó de comutação da rede ATM ao qual estão ligados utilizadores desse serviço. Designa-se por matriz de tráfego oferecido ao tráfego modelizado desta forma. Uma matriz de tráfego é uma matriz quadrada Δ cujo número de linhas (colunas) é igual ao número de nós de tráfego e em que cada elemento δ_{ij} de Δ corresponde ao tráfego oferecido do nó origem i para o nó destino j. Os elementos δ_{ii} (pertencentes à diagonal de Δ) são valores do tráfego interno ao nó i (tráfego oferecido no nó i cujo destino é o próprio nó i). O tráfego interno em cada nó deve ser considerado no dimensionamento dos nós de comutação o que, como justificado na secção 2.2.2, está fora do âmbito desta tese.

Há dois aspectos que é necessário abordar. O primeiro aspecto é como se define o tráfego entre dois nós da rede. O segundo aspecto é como se geram aleatoriamente os valores de tráfego na determinação de cenários de tráfego. No primeiro aspecto, consideramos que entre quaisquer dois nós da rede (i) a geração de VCCs é um processo de Poisson com taxa λ, (ii) a duração de cada VCC tem um tempo médio $1/\mu$ e (iii) os processos de geração e de duração dos VCCs são estatisticamente independentes. Assim, o tráfego oferecido δ_{ij} da matriz de tráfego Δ é representado por ρ_{ij} e é dado pela intensidade de tráfego em Erlangs ($\rho = \lambda / \mu$) gerado no nó i com destino ao nó j. Embora a norma ATM permita ligações ponto-a-ponto assimétricas, os serviços conversacionais actualmente existentes são simétricos pelo que as ligações estabelecidas na rede de um nó i para um nó j ocupam os mesmos recursos nos dois sentidos do trajecto por onde são estabelecidas. Neste caso, o tráfego entre 2 nós da rede é caracterizado por um valor $\rho_{\{ij\}}$ dado pela soma dos valores ρ_{ij} e ρ_{ji}.

Em relação à geração aleatória dos valores de tráfego $\rho_{\{ij\}}$ de cada serviço, a intensidade de tráfego entre dois nós da rede depende de um conjunto de factores cuja

influência não é facilmente quantificável. O número de clientes do serviço ligados aos 2 nós é, por norma, o factor mais determinante. Por exemplo, o tráfego entre Lisboa e Porto de um qualquer serviço é regra geral maior do que entre quaisquer outras duas cidades do país porque o universo de clientes é maior. No entanto, outros factores tornam esta dependência não totalmente linear tais como o interesse que os clientes de um nó têm em estabelecer ligações para o universo de clientes de outro nó (por exemplo, uma grande cidade com uma concentração de pessoas cujas origens familiares sejam de uma zona do país com menor número de habitantes faz com que o tráfego entre os nós de rede respectivos seja maior do que o número de clientes do serviço poderia à partida prever) ou as tarifas que o operador cobra para cada destino (muitos clientes reduzem ao mínimo as ligações para destinos distantes porque são mais caras). Como no caso geral o número de clientes é variável de nó para nó de rede, considera-se que a função densidade de probabilidade uniforme é uma aproximação válida para a estatística dos valores de tráfego reais $\rho_{\{ij\}}$ de qualquer serviço conversacional. Assim, o método adoptado nesta tese para a geração de cenários de tráfego baseia-se na geração aleatória de valores de tráfego com uma função densidade de probabilidade uniforme entre um valor mínimo e um valor máximo de tráfego.

O segundo caso de interesse é o tráfego uni-direccional assimétrico. Este caso corresponde aos serviços de consulta em que, para cada serviço, existem nós de tráfego e nós com servidores (não mutuamente exclusivos). As ligações são estabelecidas apenos no sentido de um nó de tráfego para um nó servidor (tráfego uni-direccional) em que no trajecto da ligação são ocupados recursos diferentes em cada um dos sentidos (tráfego assimétrico). Neste caso, define-se o tráfego oferecido à rede por cada nó origem (nó de tráfego do serviço). Ao contrário do caso anterior, em vez de uma matriz de tráfego, define-se um vector de tráfego ρ_i em que i representa um nó de tráfego do serviço. Os valores ρ_i cujo índice i representa um nó com servidor, são valores de tráfego interno ao nó i. Pelas mesmas razões apontadas para os serviços conversacionais, o método adoptado nesta tese para a geração de cenários de tráfego baseia-se na geração aleatória de valores de tráfego com uma função densidade de probabilidade uniforme entre um valor mínimo e um valor máximo de tráfego.

2.4.2 Nível da Ligação de Tráfego Ponto-Multiponto

No caso do tráfego ponto-multiponto, cada ligação submetida à rede é estabelecida de um nó origem para múltiplos nós destino por uma árvore em que são ocupados recursos em cada elemento da árvore no sentido da origem para os destinos. Embora no caso geral, por cada pedido de ligação de A para B e C seja também pedida uma ligação de B para A e C e de C para A e B, considera-se nesta tese que existe independência entre os 3 processos de geração de VCCs. Assim, para um determinado serviço ponto-multiponto, considere-se K o conjunto de todos os nós de tráfego k de um serviço e considere-se R_k^n $(1 \leq n \leq |K|-1)$ o conjunto de todas as combinações de n nós pertencente a $K \setminus \{k\}$. Considere-se também R_k a união dos conjuntos R_k^n para todos os valores de n possíveis $(1 \leq n \leq |K|-1)$, ou seja, R_k é o conjunto de todas as combinações de todos os sub-conjuntos de nós pertencentes a $K \setminus \{k\}$. Os elementos de R_k^n são designados por r_n ($r_n \in R_k^n$) e os elementos de R_k são designados por r ($r \in R_k$).

Considera-se nesta tese que de um nó origem k para um conjunto de nós destino $r \in R_k$ (i) a geração de VCCs ponto-multiponto é um processo de Poisson com taxa λ, (ii) a duração de cada VCC tem um tempo médio $1/\mu$ e (iii) os processos de geração e de duração dos VCCs são estatisticamente independentes. Assim, designa-se por ρ_k^r a intensidade de tráfego oferecida no nó k com destino para o conjunto de nós $r \in R_k$. Note-se que é possível ser definido um valor de tráfego de um nó origem para um único nó destino. Os destinos de uma ligação ponto-multiponto são clientes ligados aos nós de tráfego pelo que alguns deles podem estar ligados ao mesmo nó destino e outros podem estar ligados ao nó origem da ligação. Por exemplo, o modelo considera que uma ligação de um cliente no nó *A* para três destinos em que um cliente destino está ligado ao nó *A* e os outros dois clientes destino estão ligados ao nó *B* corresponde a ter uma ligação ponto-multiponto submetida em *A* com destino para *B* (o nó *A* faz a cópia do fluxo de células ATM da ligação estabelecida para o cliente destino que lhe está ligado e o nó *B* duplica o fluxo de células ATM para os dois clientes destino que lhe estão ligados).

Não é "realista" considerar a geração de cenários de tráfego ponto-multiponto baseada exclusivamente na geração aleatória de valores de tráfego para todos os conjuntos de destinos. De facto, há uma forte correlação entre os diferentes valores de tráfego. Por exemplo, considere-se o tráfego da origem *O* para os destinos *A*, *B* e *C*. Num cenário de

tráfego "realista", se o tráfego de *O* para *A* e o tráfego de *O* para *B* forem superiores ao tráfego de *O* para *C*, então o tráfego de *O* para *A* e *B* deverá ser também superior ao tráfego de *O* para *A* e *C* e ao tráfego de *O* para *B* e *C*. Assim, propõe-se nesta tese um método de geração de cenários de tráfego que introduz a correlação descrita. O método que se propõe nesta tese segue a seguinte abordagem:

(i) geração aleatória de valores de tráfego com uma função densidade de probabilidade uniforme entre um valor de tráfego mínimo e um valor de tráfego máximo de cada origem *k* para cada destino individual $r \in \{ R_k^n , n = 1\}$

(ii) adopção de uma fórmula que determine os valores de tráfego para múltiplos nós destino baseada nos valores gerados aleatoriamente para cada destino individual.

Neste método, os valores gerados aleatoriamente para os destinos individuais permitem reflectir o caso real em que o tráfego gerado de cada nó origem não é uniformemente distribuído pelos diferentes nós destino.

Após a geração aleatória dos valores ρ_k^d para todos os pares origem-destino (*k*,*d*) pertencentes a *K*, propõe-se que o valor de tráfego do nó origem *k* para o conjunto de nós destino r_n seja dado por:

$$\rho_k^{r_n} = \frac{\sum_{d \in r_n} \rho_k^d}{n^e \cdot p^{n-1}} \tag{8}$$

Esta fórmula determina o tráfego para um conjunto de *n* destinos através da soma dos tráfegos para cada um dos destinos individuais dividida por um factor dependente de *n* (número de nós destino) e dado pelos parâmetros *e* (inteiro) e *p* (real positivo).

É necessário agora determinar uma forma que permita escolher os parâmetros *e* e *p* de modo que os valores de tráfego ρ_k^r gerados aleatoriamente pelo método proposto tenham um conjunto de características adequado, nomeadamente um número máximo e médio de nós destino simultâneos.

Se as variáveis ρ_k^d forem geradas com média $\overline{\rho}$, então o tráfego médio total gerado de todos os nós origem para todas as combinações de *n* destinos, designado por t_n, é dado por:

$$t_n = \frac{|K| \cdot C_n^m \cdot n \cdot \overline{\rho}}{n^e \cdot p^{n-1}} \qquad m = |K| - 1 \tag{9}$$

em que $1 \le n \le m$ (C_n^m representa o número de combinações de *n* elementos num conjunto de *m* elementos). O tráfego normalizado médio, $\bar{t}_n$, é dado por:

$$\bar{t}_n = \frac{t_n}{\sum_{i=1}^{m} t_n} = \frac{C_n^m \Big/ \left(n^{e-1} \cdot p^{n-1}\right)}{\sum_{i=1}^{m} C_i^m \Big/ \left(i^{e-1} \cdot p^{i-1}\right)} \qquad m = |K| - 1 \tag{10}$$

Note-se que a equação (10) não depende da média ($\overline{\rho}$) dos valores gerados aleatoriamente. O conjunto de valores $\bar{t}_n$ define o perfil de tráfego gerado pelo método proposto. A Figura 2-3 ilustra o perfil de tráfego para diferentes valores de *e* e *p* no caso em que |*K*| = 30. Pode-se verificar que quando se aumentam os valores de *e* ou de *p*, faz-se com que o tráfego para um número maior de destinos simultâneos fique percentualmente mais pequeno. Mediante uma correcta parametrização dos valores de *e* e *p*, é possível gerar cenários de tráfego com diferentes perfis de tráfego.

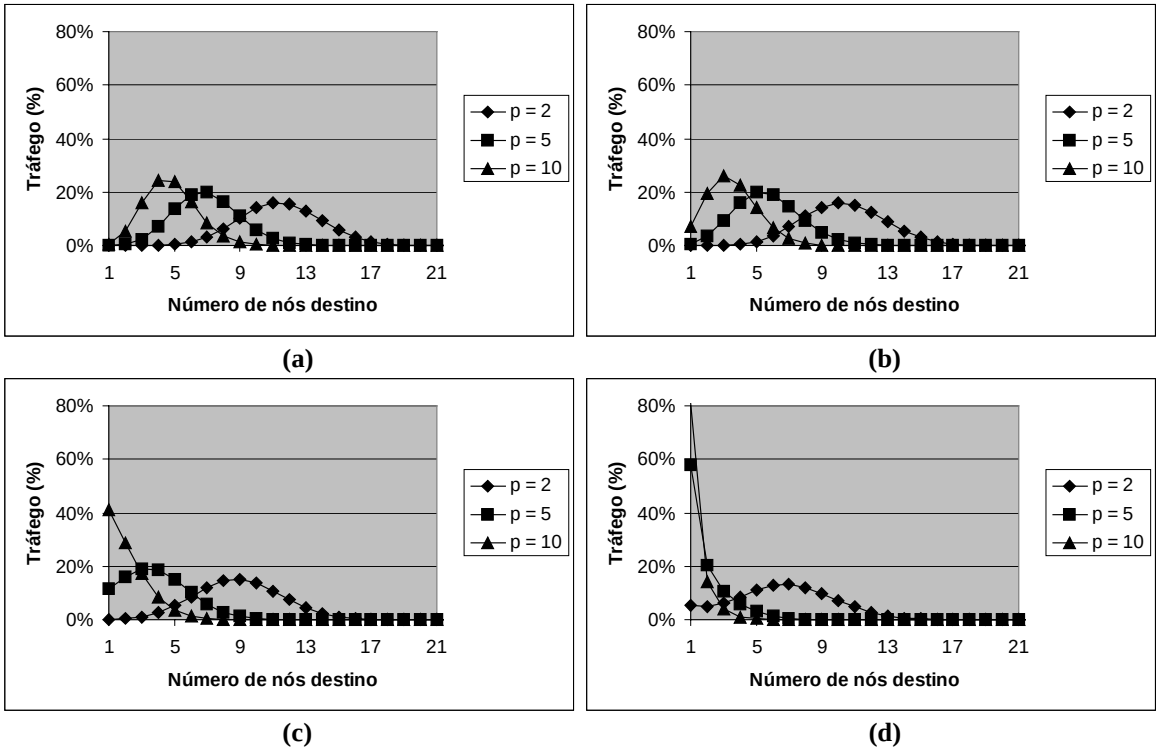


Figura 2-3 : Tráfego normalizado médio para um universo de 30 nós de tráfego
(a) *e* = -2 (b) *e* = 0 (c) *e* = 2 (d) *e* = 4

Um perfil de tráfego é caracterizado por duas grandezas significativas: o número máximo e número médio de nós destino simultâneos.

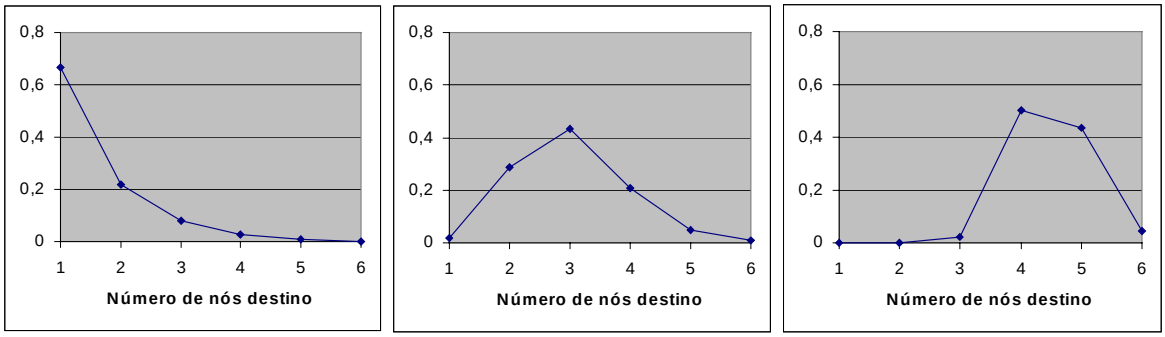
Na prática, o número máximo de participantes numa ligação ponto-multiponto é menor do que o número total de nós de tráfego do serviço pelo que se considera que um serviço baseado em ligações ponto-multiponto tem valores de tráfego desprezáveis para ligações com muitos destinos simultâneos. Para além disso, como alguns terminais destino podem pertencer ao mesmo nó destino, o número máximo de nós destino simultâneos torna-se ainda menor em termos médios. Assim, designando o número máximo de nós destino por *b* e a percentagem de tráfego desprezado por *q*, queremos escolher um perfil de tráfego em que a soma dos valores de $\bar{t}_n$, para *n* maior do que *b*, seja menor ou igual que a percentagem de tráfego desprezado *q*. Esta relação é dada pela seguinte equação:

$$\frac{\sum_{n=b+1}^{m}\left(\frac{C_n^m}{n^{e-1}\cdot p^{n-1}}\right)}{\sum_{n=1}^{m}\left(\frac{C_n^m}{n^{e-1}\cdot p^{n-1}}\right)} \le q \qquad (11)$$

O número médio de nós destino de um cenário de tráfego define o grau de diversidade do tráfego submetido à rede. Quanto maior for este valor, maior é a percentagem de tráfego para múltiplos destinos e maiores podem ser os ganhos de utilização de recursos que se obtêm ao se explorar a capacidade dos comutadores ATM poderem copiar fluxos de células de entrada para múltiplas portas de saída. O número médio de nós destino é dado pela média pesada dos valores de *n* entre 1 e *b* em que os pesos são dados pelos valores $\bar{t}_n$. A relação entre o número médio de nós destino, designado por *g*, e os valores de tráfego normalizado médio é dada por:

$$g = \frac{\sum_{n=1}^{b}\left(n\cdot\frac{C_n^m}{n^{e-1}\cdot p^{n-1}}\right)}{\sum_{n=1}^{b}\left(\frac{C_n^m}{n^{e-1}\cdot p^{n-1}}\right)} \qquad (12)$$

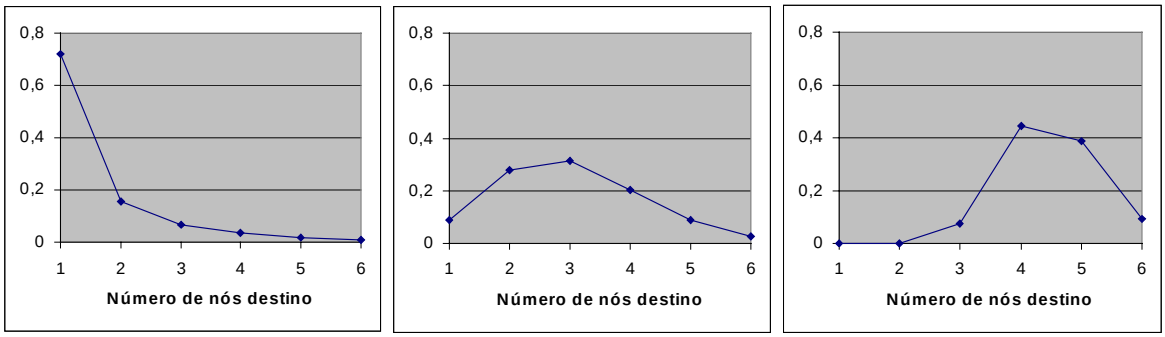
As equações (11) e (12) permitem determinar os parâmetros *e* e *p* em função do número total de nós de tráfego |*K*|, (*m* = |*K*| − 1), do número máximo de nós destino *b*, do número médio de nós destino *g* e da percentagem de tráfego desprezada *q*. No Anexo A é apresentado um método baseado em análise numérica capaz de resolver estas duas equações.



(a) **(b)** **(c)**

Figura 2-4 : Perfis de tráfego para $|K| = 30$, $b = 6$ e $q = 0{,}1\%$

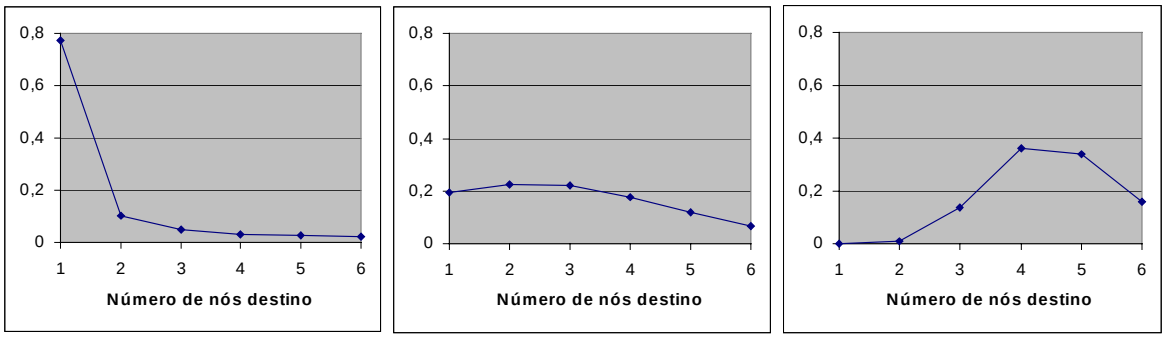
(a) $g = 1.5$, (b) $g = 3.0$, (c) $g = 4.5$



(a) **(b)** **(c)**

Figura 2-5 : Perfis de tráfego para $|K| = 30$, $b = 6$ e $q = 1\%$

(a) $g = 1.5$, (b) $g = 3.0$, (c) $g = 4.5$



(a) **(b)** **(c)**

Figura 2-6 : Perfis de tráfego para $|K| = 30$, $b = 6$ e $q = 5\%$

(a) $g = 1.5$, (b) $g = 3.0$, (c) $g = 4.5$

Considerando 30 nós de tráfego e um máximo de 6 nós destino, os perfis de tráfego resultantes das equações (11) e (12) para 3 valores de número médio de destinos são apresentados na Figura 2-4 (tráfego desprezado de 0,1%), Figura 2-5 (tráfego desprezado de 1%) e Figura 2-6 (tráfego desprezado de 5%). As figuras ilustram que quanto maior é o valor médio de número de destinos, maior é a percentagem de tráfego para um número

maior de destinos simultâneos. A comparação das diferentes figuras permite verificar o efeito de diferentes valores de tráfego desprezado nos respectivos perfis. Um valor de tráfego desprezado menor faz com que o perfil tenha um pico maior para os valores de *n* em torno da média (diminuindo assim a variância do perfil). Este efeito é maior para valores grandes de número médio de nós destino. A inexistência de dados experimentais que caracterizem o processo de geração de VCCs ponto-multiponto não nos permite adoptar nenhum critério fundamentado em relação ao valor de tráfego desprezado mais adequado para cada perfil. Assim, consideramos um valor de q = 1% para a geração de cenários de tráfego nos capítulos seguintes.

2.4.3 Nível da Célula

Após um VCC ser estabelecido numa rede lógica, o emissor envia células com características previamente contratadas e a rede assegura um conjunto de parâmetros de qualidade de serviço (também previamente contratados) na transmissão das células até ao destino (ou destinos no caso de ligações ponto-multiponto). Ao nível da célula, o modelo de cada VPC é uma linha de transmissão, com uma capacidade especificada em bits por segundo, em que à sua entrada existe uma fila de espera com uma disciplina de atendimento FIFO (*first in, first out*). Assim, numa ligação lógica (ou VCC), as células percorrem tantos sistemas destes quantos os VPCs pelos quais ela é estabelecida. Ao passarem pelas filas de espera, as células podem ser descartadas se a fila estiver cheia. A largura de banda que é preciso reservar para um VCC em cada um dos VPCs por onde ela atravessa é dependente da estatística de geração de células do VCC, da estatística de geração de células dos outros VCCs estabelecidos pelos mesmos VPCs, da capacidade de cada VPC e dos parâmetros máximos de perda, de atraso e de variação do atraso das células.

No caso de VCCs cujo perfil de tráfego é constante (geração de células a um ritmo constante no tempo), a largura de banda a reservar em cada um dos VPCs é igual à largura de banda da fonte. No caso de ligações cujo perfil de tráfego é variável com o tempo, é teoricamente possível reservar uma largura de banda em cada VPC menor do que o ritmo de pico da geração de células do VCC (característica inerente à multiplexagem estatística nas filas de espera) [Anick82] [Elawlid93] [Gibbens91] [Kesidis93] [Kelly91] [Yang95]. No caso geral, a largura de banda a reservar por ligação depende do número máximo de ligações a suportar em cada uma das fila de espera: quanto maior o número de ligações,

mais ganhos estatísticos se obtêm e menor é o valor da largura de banda necessária por ligação. No entanto, existem alguns métodos que recorrem à noção de largura de banda efectiva, um valor único de capacidade a reservar em todos os VPCs que não depende do número máximo de ligações simultâneas a suportar por cada VPC [Elawlid93] [Gibbens91] [Kesidis93] [Kelly91]. Estes métodos, regra geral, são mais conservativos que os anteriores no sentido que determinam valores de capacidade maiores para os mesmos modelos da fonte de células mas permitem o cálculo das capacidades de redes lógicas em que as ligações atravessam múltiplos VPCs [Veciana94]. Os modelos de dimensionamento apresentados nesta tese assumem que as ligações dos serviços suportados são caracterizadas por um único valor de largura de banda determinado por um método de largura de banda efectiva, se for conhecido, ou considerando o valor de pico das ligações.

2.5 Definição de Probabilidade de Bloqueio

O parâmetro mais importante na definição da qualidade de serviço ao nível da ligação é a probabilidade de bloqueio. Este parâmetro representa quantas vezes (em percentagem) é que em média o pedido de estabelecimento de um VCC pode ser recusado por falta de recursos na rede. No caso de ligações ponto-a-ponto (serviços conversacionais e serviços de consulta), quando um VCC é submetido à rede, o VCC é estabelecido desde que haja recursos disponíveis no trajecto definido da origem para o destino. Assim, a decisão é sim ou não e a probabilidade de bloqueio para cada classe de tráfego é dada pela percentagem de VCCs submetidos à rede que são recusados por não haver recursos disponíveis [COST94] [Kelly86]. O caso de VCCs ponto-multiponto é mais complexo pois pode acontecer que não seja possível estabelecer o VCC apenas para um sub-conjunto dos destinos. Neste caso, a rede ATM estabelece o VCC com os destinos para os quais existem recursos na rede. A definição da probabilidade de bloqueio do tráfego deve ter em consideração este facto. Considere-se o seguinte exemplo: um cliente ligado ao nó *A* pede o estabelecimento de um VCC ponto-multiponto para um cliente ligado ao nó *B* e um cliente ligado ao nó *C*. Se o VCC for estabelecido apenas para *B*, o bloqueio pode ser considerado de duas formas alternativas:

- um bloqueio de 50% pois a ligação foi estabelecida para metade dos destinos;
- um bloqueio de 0% para o destino *B* e um bloqueio de 100% para o destino C.

No primeiro caso, a probabilidade de bloqueio B_f é definida para a classe de tráfego *f* cuja origem é A e cujo conjunto destino é {B,C}. No segundo caso, a probabilidade de bloqueio B_o^d é definida para cada par origem-destino. Considere-se o seguinte exemplo das ligações submetidas à rede do nó origem *A* para os nós destino *B* e *C*. Considere o número de VCCs submetidos à rede, durante um período de tempo longo, e destes, o número dos que foram bloqueados, dado pela seguinte tabela:

Classe de tráfego	**Origem**	**Destino(s)**	**Ligações submetidas**	**Ligações bloqueadas**
1	*A*	*B*	50	2
2	*A*	*C*	100	3
3	*A*	*B* e *C*	80	3 para *B* 5 para *C* 2 para *B* e *C*

No caso de probabilidade de bloqueio por classe de tráfego B_f, temos que:

$$B_1 = \frac{2}{50} = 4.0\%$$

$$B_2 = \frac{3}{100} = 3.0\%$$

$$B_3 = \frac{3\times 50\% + 5\times 50\% + 2\times 100\%}{80} = 7.5\%$$

No caso de probabilidades de bloqueio por par origem-destino B_o^d, temos que:

$$B_A^B = \frac{2+3+2}{50+80} \approx 5.4\%$$

$$B_A^C = \frac{3+5+2}{100+80} \approx 5.6\%$$

Neste último caso, por exemplo, o denominador é o número total de VCCs em que o nó *C* é destino e o numerador é o número de VCCs em que o destino *C* foi bloqueado. Esta definição quantifica a probabilidade de bloqueio para cada destino independentemente da ligação se destinar a outros nós da rede.

Embora as duas alternativas possam ser válidas como definição de probabilidade de bloqueio para tráfego ponto-multiponto, consideramos que a segunda definição reflecte melhor a noção de bloqueio que o utilizador sente perante o serviço que lhe é prestado. Por exemplo, se um comerciante estabelece periodicamente uma ligação ponto-multiponto de

video-conferência com quatro fornecedores em que um está no Porto, outro em Lisboa, outro em Coimbra e outro em Aveiro e se a maior parte das vezes consegue estabelecer a ligação para o Porto, Coimbra e Aveiro e não consegue para Lisboa, o comerciante tem tendência para pensar que tem um mau serviço para Lisboa (alta probabilidade de bloqueio com este destino) em vez de pensar que tem um serviço com uma taxa de bloqueio de 25% (1 destino bloqueado em 4 destinos tentados). Assim, o trabalho apresentado nesta tese adopta a segunda definição para caracterizar a probabilidade de bloqueio de um serviço ponto-multiponto:

"A probabilidade de bloqueio de um nó origem *o* para um nó destino *d* é a probabilidade de uma chamada ponto-multiponto de *o* para qualquer conjunto de nós destino que inclua *d* não ser estabelecida para *d*."

2.6 Problema de Dimensionamento: Integralidade dos VPCs

Esta secção tem dois objectivos. O primeiro é ilustrar de que forma se podem combinar os modelos elementares introduzidos neste capítulo na formulação de um problema de dimensionamento da rede. O segundo objectivo é introduzir a questão da integralidade dos VPCs e a discussão da utilidade de modelos de dimensionamento que a contemplem.

Quando as redes lógicas de suporte a cada serviço são constituídas apenas por VPCs de extremos definidos, a topologia da rede lógica é conhecida à partida e as capacidades de cada VPC podem ser calculadas independentemente dos seus trajectos. Considere-se então o caso em que todos os VPCs têm extremos definidos e são simétricos com capacidades b_k^s conhecidas. O problema de dimensionamento da rede pode ser enunciado com base na notação introduzida neste capítulo. Assim, considere-se que a rede é definida por:

N Conjunto de nós da rede

A Conjunto de arestas $\{i,j\}$ da rede: (i,j) representa a aresta $\{i,j\}$ no sentido de i para j e (j,i) representa o sentido inverso

em que as ligações físicas que podem ser implementados em cada uma das arestas são caracterizadas por:

T Conjunto dos tipos t de ligações físicas possíveis de implementar entre os nós de rede

α_t Largura de banda de uma ligação física do tipo t

$Y_{\{ij\}}^t$ Número máximo de ligações físicas do tipo t entre os nós i e j

$C_{\{ij\}}^t$ Custo associado à utilização de uma ligação do tipo t entre os nós i e j

e na qual se pretende configurar um conjunto de redes lógicas de suporte caracterizadas por VPCs simétricos de extremos definidos :

S Conjunto de serviços s suportados pela rede

K_s Conjunto de VPCs k da rede lógica de suporte ao serviço s; cada VPC k é caracterizado por um nó origem (o_k^s), um nó destino (d_k^s) e uma capacidade (b_k^s)

Considerem-se as variáveis inteiras não negativas $y_{\{ij\}}^t$ que determinam o número de ligações físicas do tipo t entre os nós i e j. Usando a modelização explícita de VPCs, considerem-se as variáveis binárias x_{ij}^{sk} que, quando têm o valor 1, significam que o VPC k de suporte ao serviço s passa no arco (i,j). Com a notação apresentada, é possível enunciar o problema de dimensiomento da rede através do seguinte modelo de Programação Inteira:

Minimizar $$\sum_{\{i,j\}\in A}\sum_{t\in T} C_{\{ij\}}^t \cdot y_{\{ij\}}^t \qquad (13a)$$

Sujeito a:

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{ij}^{sk} - x_{ji}^{sk}\right)=1 \qquad , s\in S\,,\, k\in K_s\,,\, i = o_k^s \qquad (13b)$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{ji}^{sk} - x_{ij}^{sk}\right)=1 \qquad , s\in S\,,\, k\in K_s\,,\, i = d_k^s \qquad (13c)$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{ij}^{sk} - x_{ji}^{sk}\right)=0 \qquad , s\in S\,,\, k\in K_s\,,\, i\in N\setminus\{o_k^s, d_k^s\} \qquad (13d)$$

$$\sum_{s\in S}\sum_{k\in K_s} b_k^s\left(x_{ij}^{sk} + x_{ji}^{sk}\right) \le \sum_{t\in T}\alpha_t \cdot y_{\{ij\}}^t \qquad , \{i,j\}\in A \qquad (13e)$$

$$y_{\{ij\}}^t \le Y_{\{ij\}}^t \qquad , \{i,j\}\in A\,,\, t\in T \qquad (13f)$$

$x_{ij}^{sk} \in \{0,1\}$ $y_{\{ij\}}^t$ inteiro não negativo

A função objectivo é a função de custo da rede, igual à expressão (4). Como no caso apresentado os VPCs são todos com extremos definidos, as equações (13b), (13c) e (13d) são incluídas nas restrições do problema na mesma forma das equações (5) apresentadas anteriormente. A restrição (13e) garante que a capacidade total das ligações físicas entre dois nós de rede, expressão (1), é maior ou igual do que a soma das larguras de banda dos VPCs que passam pelas ligações.

A restrição (13e) é uma aproximação do problema real porque assume que a capacidade de um VPC pode ser distribuída por duas (ou mais) ligações físicas em paralelo. Os equipamentos actualmente existentes não permitem segmentar um VPC por diferentes ligações físicas. Embora existam normas recentes [ATMForum99] que permitem considerar como uma única ligação a agregação das capacidades de diferentes ligações físicas entre 2 comutadores, estas normas aplicam-se para já apenas a ligações físicas de pequena capacidade tais como o DS1 (1.5 Mbps) e E1 (2 Mbps). Assim, no caso geral isto não é possível e a capacidade de cada VPC que atravessa o conjunto de ligações físicas entre dois nós de rede deve ser suportada integralmente por uma das ligações físicas: esta é a propriedade da integralidade dos VPCs.

Uma forma possível de formular o problema de dimensionamento anterior tendo em conta a integralidade dos VPCs, é extendendo cada uma das variáveis $y^{t}_{\{ij\}}$ inteiras num conjunto de variáveis binárias $y^{tn}_{\{ij\}}$ que determinam se a *n*-ésima interface do tipo *t* é usada entre os nós *i* e *j*. Da mesma forma as variáveis x^{sk}_{ij} deverão ser extendidas para x^{sktn}_{ij} de modo a indicarem para cada arco (*i*,*j*), em que ligação física passam. O modelo de dimensionamento resultante é dado por:

Minimizar $$\sum_{\{i,j\}\in A}\sum_{t\in T} C^{t}_{\{ij\}} \sum_{n=1}^{Y^{t}_{\{ij\}}} y^{tn}_{\{ij\}} \qquad (14a)$$

Sujeito a:

$$\sum_{j\in N:\{i,j\}\in A}\sum_{t\in T}\sum_{n=1}^{Y^{t}_{\{ij\}}}\left(x^{sktn}_{ij} - x^{sktn}_{ji}\right)=1 \qquad , s\in S\ ,\ k\in K_s\ ,\ i = o^{s}_{k} \qquad (14b)$$

$$\sum_{j\in N:\{i,j\}\in A}\sum_{t\in T}\sum_{n=1}^{Y^{t}_{\{ij\}}}\left(x^{sktn}_{ji} - x^{sktn}_{ij}\right)=1 \qquad , s\in S\ ,\ k\in K_s\ ,\ i = d^{s}_{k} \qquad (14c)$$

$$\sum_{j \in N : \{i,j\} \in A} \sum_{t \in T} \sum_{n=1}^{Y_{\{ij\}}^t} \left(x_{ij}^{sktn} - x_{ji}^{sktn} \right) = 0 \qquad , s \in S , k \in K_s , i \in N \setminus \{ o_k^s , d_k^s \} \qquad (14d)$$

$$\sum_{s \in S} \sum_{k \in K_s} b_k^s \left(x_{ij}^{sktn} + x_{ji}^{sktn} \right) \leq \alpha_t \cdot y_{\{ij\}}^{tn} \qquad , \{i,j\} \in A , t \in T , 1 \leq n \leq Y_{\{ij\}}^t \qquad (14e)$$

$$x_{ij}^{sktn} \in \{0,1\} \quad y_{\{i,j\}}^{tn} \in \{0,1\}$$

As restrições (14b), (14c) e (14d) são as restrições típicas de VPCs com extremos definidos. A restrição (14e) assegura a integralidade dos VPC pois ela garante que cada ligação física tem uma largura de banda maior ou igual do que os VPCs que por ela passam.

No entanto, este modelo é muito mais complexo que o anterior. Se considerarmos que $Y_{\{ij\}}^t$ é igual a Y para todo o t e todas as arestas $\{i,j\}$, então neste modelo o número de variáveis y é Y vezes maior e o número de variáveis x é $|T| \times Y$ maior. O número de restrições (14e) é também $|T| \times Y$ vezes maior. Isto significa que o problema de dimensionamento é mais complexo e necessariamente mais difícil de resolver. Esta é a razão pela qual o presente trabalho não contempla a integralidade dos VPCs. Não há conhecimento na literatura científica que esta questão tenha alguma vez sido abordada. Uma rede dimensionada sem considerar a integralidade dos VPCs assume um pós-processamento da solução encontrada. Este pós-processamento consiste em verificar para cada aresta se é possível condicionar os VPCs no conjunto de ligações físicas determinadas e, se não for, encontrar uma outra combinação de ligações físicas de menor custo que os consiga condicionar.

CAPÍTULO 3

REDES LÓGICAS COM VPCs EXTREMO-A-EXTREMO

Este capítulo aborda o problema de dimensionamento de redes ATM em que cada serviço é suportado por uma rede lógica constituída apenas por VPCs extremo-a-extremo. A adopção de redes lógicas com esta arquitectura tem como principal vantagem a simplificação dos mecanismos de CAC e, consequentemente, a minimização dos tempos de estabelecimento de VCCs. Um VPC extremo-a-extremo suporta apenas a classe de tráfego entre os seus nós extremos. Neste caso, é possível separar o cálculo da capacidade dos VPCs do problema de dimensionamento da rede física que inclui a determinação dos trajectos dos VPCs.

Neste capítulo é também abordado o encaminhamento multi-horário. No dimensionamento multi-horário, assume-se que os VPCs podem ser reconfigurados ao longo do tempo para acompanharem variações do tráfego oferecido à rede. Neste caso, o tempo é particionado em intervalos de tempo definidos e o trajecto e a capacidade de cada VPC podem variar na transição entre intervalos de tempo.

Embora os modelos de dimensionamento propostos sejam modelos de Programação Inteira, não é conhecido nenhum algoritmo que permita calcular em tempo útil as soluções óptimas para problemas razoavelmente grandes. Assim, neste capítulo é proposto um algoritmo heurístico baseado numa técnica da Investigação Operacional conhecida por

relaxação Lagrangeana com optimização por sub-gradiente para a determinação de boas soluções no dimensionamento de redes ATM.

Este capítulo está organizado da seguinte forma. A secção 3.1 endereça o problema do cálculo da capacidade dos VPCs para os três tipos de serviços abordados: serviços conversacionais, serviços de consulta e serviços ponto-multiponto. Na secção 3.2 são apresentados dois modelos uni-horário de Programação Inteira cuja resolução permite dimensionar redes sem capacidade de reconfiguração dinâmica de VPCs: um modelo baseia-se na modelização explícita de VPCs e o outro modelo baseia-se na modelização por trajectos candidatos. O dimensionamento de redes com reconfiguração dinâmica é abordado na secção 3.3 onde são apresentados os modelos multi-horário adequados. Na secção 3.4 é proposto um algoritmo heurístico para a resolução dos modelos apresentados baseado na relaxação Lagrangeana com optimização por sub-gradiente. Finalmente, na secção 3.5 são apresentados resultados computacionais do algoritmo proposto em que se afere da utilizabilidade do algoritmo bem como da qualidade das soluções por ele encontradas.

3.1 Cálculo da Capacidade dos VPCs

Conforme os serviços suportados, os VPCs podem ser simétricos ou assimétricos. Daí a necessidade de para cada VPC se definir um nó origem e um nó destino. Desta forma é possível designar a capacidade do VPC nos seus dois sentidos: b_k^s designa a capacidade da origem para o destino e $\underline{b}_k^s$ designa a capacidade do destino para a origem. Quanto à definição dos extremos de cada VPC, existem também dois casos possíveis: VPCs com extremos definidos e VPCs com destino indefinido (este último caso para os serviços de consulta). Esta secção endereça o problema do cálculo da capacidade de cada VPC (valores de b_k^s e $\underline{b}_k^s$). Este problema é diferente para cada tipo de serviço pelo que é abordado separadamente na secção 3.1.1 para os serviços conversacionais, na secção 3.1.2 para os serviços de consulta e na secção 3.1.3 para os serviços ponto-multiponto.

3.1.1 Serviços Conversacionais

Para serviços conversacionais simétricos, uma rede lógica com VPCs extremo-a-extremo inclui um VPC para cada par de nós de tráfego. Assim, considere-se que entre os

nós de tráfego *i* e *j* do serviço $s \in S$, existe o VPC $k \in K_s$ de suporte a uma intensidade de tráfego ρ_k. Considere também que os VCCs pertencentes ao serviço *s* ocupam uma largura de banda w_s. O cálculo da capacidade b_k^s (note-se que se considera os serviços simétricos pelo que $\underline{b}_k^s = b_k^s$) do VPC *k* recorre à fórmula de ErlangB e é enunciado pelo seguinte problema de minimização:

Minimizar m_k^s

Sujeito a

$$ER\left[\rho_k, m_k^s\right] \leq P_s$$

em que m_k^s é uma variável inteira não negativa que representa a capacidade do VPC *k* em número de VCCs e P_s representa a probabilidade de bloqueio máxima para os VCCs do serviço *s*. $ER[\rho,m]$ representa a fórmula de ErlangB:

$$ER[\rho, m] = \frac{\rho^m / m!}{\sum_{i=0}^{m} \rho^i / i!}$$

e determina a probabilidade de bloqueio de cada ligação de um sistema ponto-a-ponto quando o tráfego de chegada é ρ e o sistema tem capacidade para *m* ligações em simultâneo. Dado m_k^s, o valor de b_k^s é dado por:

$$b_k^s = w_s \cdot m_k^s$$

É possível provar que a fórmula de ErlangB $ER[\rho,m]$ é decrescente com *m* para um valor fixo de ρ. Assim, o problema de minimização enunciado é de fácil resolução: percorrem-se os valores de *m* não negativos começando no valor 1 e terminando no primeiro valor para o qual a restrição do problema seja válida.

3.1.2 Serviços de Consulta

Num serviço de consulta, os VCCs são estabelecidos de um nó de tráfego para um nó servidor pelo que uma rede lógica de suporte inclui um VPC entre cada nó de tráfego e um qualquer dos nós servidores. Não existe tráfego oferecido de um nó servidor para um nó de tráfego. Por convenção, considera-se o nó de tráfego como o nó origem e o nó servidor como o nó destino de cada VPC. Os VCCs pertencentes a um serviço *s* de consulta são agora caracterizadas por ocuparem uma largura de banda w_s no sentido da origem para o

destino e uma largura de banda $\underline{w}_s$ no sentido do destino para a origem. Apesar das ligações serem assimétricas, o tráfego suportado por cada VPC é submetido apenas num dos seus extremos (por convenção, na origem) pelo que o cálculo das capacidades em número de VCCs do VPC k (m_k^s) segue o mesmo método que foi apresentado para os serviços de consulta. Dado o valor de m_k^s, os valores de b_k^s e $\underline{b}_k^s$ são dados por:

$$b_k^s = w_s \cdot m_k^s$$
$$\underline{b}_k^s = \underline{w}_s \cdot m_k^s$$

3.1.3 Serviços Ponto-Multiponto

Tal como para serviços conversacionais, para os serviços ponto-multiponto uma rede lógica com VPCs extremo-a-extremo inclui um VPC para cada par de nós de tráfego. Numa ligação ponto-multiponto, o nó origem é o responsável por copiar as células do VCC para os VPCs que o ligam aos nós destino. Assim, ao contrário dos serviços conversacionais, o tráfego oferecido a um VPC é diferente em cada um dos seus sentidos e é uma soma de vários valores de tráfego. A Figura 3-1 ilustra a forma como o tráfego oferecido é encaminhado por uma rede lógica de VPCs extremo-a-extremo. Este exemplo contempla três nós de tráfego A, B e C. O VPC entre os nós A e B suporta no sentido de A para B o tráfego de A para B (ρ_A^B) e o tráfego de A para B e C simultaneamente (ρ_A^{BC}). No sentido inverso, este VPC suporta o tráfego de B para A (ρ_B^A) e o tráfego de B para A e C simultaneamente (ρ_B^{AC}).

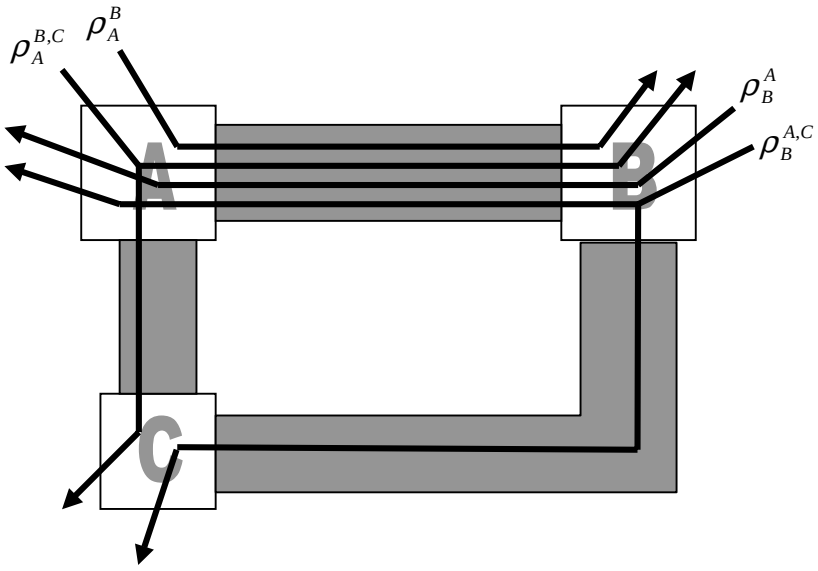


Figura 3-1 : Tráfego ponto-multiponto oferecido a VPCs extremo-a-extremo

No caso geral, um VPC *k* de suporte a um serviço ponto-multiponto *s* entre os nós *i* e *j* suporta no sentido de *i* para *j* todo o tráfego de *i* para as combinações de destinos que contenham o nó *j*. Assim, se por convenção o nó *i* for a origem e o nó *j* for o destino, então o tráfego suportado pelo VPC *k* pertencente à rede de suporte do serviço *s* no sentido da origem para o destino (ρ_k^s) é dado por:

$$\rho_k^s = \sum_{r \in R_i^s : j \in r} \rho_i^r \qquad \text{, } i \text{ é a origem de } k\text{, } j \text{ é o destino de } k$$

em que R_i^s é o conjunto das combinações *r* de nós para as quais existe tráfego do serviço *s* oferecido no nó *i*. No sentido contrário, o tráfego suportado pelo VPC *k* pertencente à rede de suporte do serviço *s* ($\underline{\rho}_k^s$) é dado por:

$$\underline{\rho}_k^s = \sum_{r \in R_j^s : i \in r} \rho_j^r \qquad \text{, } i \text{ é a origem de } k\text{, } j \text{ é o destino de } k$$

Relembre-se que, conforme discutido na secção 2.5, a probabilidade de bloqueio de um nó origem *o* para um nó destino *d* é a probabilidade de uma chamada ponto-multiponto de *o* para qualquer conjunto de nós destino que inclua *d* não ser estabelecida para *d*. No caso de VPCs extremo-a-extremo, a probabilidade de bloqueio de um nó origem *o* para um nó destino *d* é a probabilidade de bloqueio do VPC que os liga. Como os valores de tráfego submetido a um VPC são diferentes em cada um dos seus sentidos, então a capacidade do VPC *k* de um serviço *s* em número de ligações simultâneas deverá ser calculada separadamente para os dois sentidos (m_k^s e $\underline{m}_k^s$) com base nos valores de tráfego (ρ_k^s e $\underline{\rho}_k^s$). Cada um destes valores é calculado pelo mesmo método que foi apresentado para os serviços de consulta. Considerando-se que as ligações pertencentes ao serviço *s* ocupam uma largura de banda w_s em cada uma das ligações da árvore, os valores de b_k^s e $\underline{b}_k^s$ são então dados por:

$$b_k^s = w_s \cdot m_k^s$$
$$\underline{b}_k^s = w_s \cdot \underline{m}_k^s$$

3.2 Modelo de Dimensionamento Uni-Horário

Uma forma de definir os parâmetros associados aos VPCs de todas as redes lógicas que seja genérica para qualquer um dos tipos de serviços endereçados é a seguinte:

S Conjunto de serviços *s* suportados pela rede

K_s Conjunto de VPCs *k* da rede lógica de suporte ao serviço *s*; cada VPC *k* é caracterizado por:

- o_k^s nó origem
- D_k^s conjunto de possíveis nós destinos
- b_k^s capacidade no sentido da origem para o destino
- $\underline{b}_k^s$ capacidade no sentido do destino para a origem

Note-se que para os VPCs de suporte aos serviços conversacionais b_k^s é igual a $\underline{b}_k^s$ e que para os VPCs de suporte aos serviços conversacionais e serviços ponto-multiponto, D_k^s é um conjunto singular.

Nas sub-secções seguintes propõem-se dois modelos de Programação Inteira cuja resolução permite, com base nos valores de capacidade de cada VPC calculados previamente, dimensionar a rede física de menor custo e simultaneamente determinar uma solução para os trajectos de todos os VPCs na rede física resultante. Para os VPCs de suporte a serviços de consulta, a resolução do modelo determina também qual o nó servidor a que o VPC de cada nó de tráfego deve ser ligado.

O primeiro modelo proposto baseia-se na modelização explícita de VPCs apresentada na secção 2.3.1. O segundo modelo proposto baseia-se na modelização de VPCs por trajectos candidatos apresentada na secção 2.3.2. No segundo modelo, para além das capacidades de cada VPC, é assumido também que os trajectos candidatos são determinados à priori.

No caso dos VPCs com destino indefinido (os VPCs de suporte a serviços de consulta), assume-se que, qualquer que seja a forma de determinar os trajectos candidatos, são excluidos os trajectos que para atingir um determinado nó servidor passem por outro nó servidor. A razão da restrição enunciada para a determinação dos trajectos candidatos reside no seguinte argumento: se o VPC que liga um nó de tráfego atinge um nó servidor, considerar que ele continua até outro nó servidor resulta necessariamente numa ocupação extra de capacidade nas ligações físicas e, consequentemente, numa solução de rede que é de custo igual ou superior. Assim, considerar estes trajectos como candidatos apenas aumenta a complexidade do modelo sem que a solução do problema assim formulado possa ser melhor.

3.2.1 Formulação por Modelização Explícita de VPCs

N Conjunto de nós da rede

A Conjunto de arestas $\{i,j\}$ da rede: (i,j) representa a aresta $\{i,j\}$ no sentido de i para j e (j,i) representa o sentido inverso

T Conjunto dos tipos t de ligações físicas possíveis de implementar entre os nós de rede

α_t Largura de banda de uma ligação física do tipo t

$Y_{\{ij\}}^t$ Número máximo de ligações físicas do tipo t entre os nós i e j

$C_{\{ij\}}^t$ Custo associado à utilização de uma ligação do tipo t entre os nós i e j

S Conjunto de serviços s suportados pela rede

K_s Conjunto de VPCs k da rede lógica de suporte ao serviço s; cada VPC k é caracterizado por:

- o_k^s Nó origem
- D_k^s Conjunto de possíveis nós destinos
- b_k^s Largura de banda no sentido da origem para o destino
- $\underline{b}_k^s$ Largura de banda no sentido do destino para a origem

$y_{\{ij\}}^t$ Variável inteira que determina o número de ligações físicas do tipo t entre os nós i e j

x_{ij}^{sk} Variável binária de trajecto: 1 significa que o VPC k de suporte ao serviço s passa no arco (i,j)

Minimizar $$\sum_{\{i,j\}\in A}\sum_{t\in T} C_{\{ij\}}^t \cdot y_{\{ij\}}^t \qquad (1a)$$

Sujeito a:

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{ij}^{sk} - x_{ji}^{sk}\right)=1 \qquad , s\in S\,,\, k\in K_s\,,\, i = o_k^s \qquad (1b)$$

$$\sum_{i\in D_k^s}\sum_{j\in N:\{i,j\}\in A}\left(x_{ji}^{sk} - x_{ij}^{sk}\right)=1 \qquad , s\in S\,,\, k\in K_s \qquad (1c)$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{ij}^{sk} - x_{ji}^{sk}\right)=0 \qquad , s\in S\,,\, k\in K_s\,,\, i\in N\setminus\{o_k^s, D_k^s\} \qquad (1d)$$

$$\sum_{s\in S}\sum_{k\in K_s}\left(b_k^s\cdot x_{ij}^{sk} + \underline{b}_k^s\cdot x_{ji}^{sk}\right) \le \sum_{t\in T}\alpha_t\cdot y_{\{ij\}}^t \qquad , (i,j)\in A \qquad (1e)$$

$$y_{\{ij\}}^t \le Y_{\{ij\}}^t \qquad , \{i,j\}\in A\,,\, t\in T \qquad (1f)$$

$x_{ij}^{sk} \in \{0,1\}$ $y_{\{ij\}}^t$ inteiro não negativo

As restrições (1b), (1c) e (1d) são as restrições introduzidas na secção 2.3.1 para VPCs com destino indefinido. As restrições (1e) garantem que a capacidade instalada em cada aresta é suficiente para suportar todos os VPCs que por lá passam. Note-se que como os serviços podem ser assimétricos é necessário que a capacidade instalada seja suficiente em ambos os sentidos da aresta pelo que as restrições (1e) aplicam-se aos dois arcos de cada aresta. Finalmente, as restrições (1f) garantem que o número máximo de ligações físicas não é ultrapassado na solução encontrada.

3.2.2 Formulação por Trajectos Candidatos

N Conjunto de nós da rede

A Conjunto de arestas $\{i,j\}$ da rede: (i,j) representa a aresta $\{i,j\}$ no sentido de i para j e (j,i) representa o sentido inverso

T Conjunto dos tipos t de ligações físicas possíveis de implementar entre os nós de rede

α_t Largura de banda de uma ligação física do tipo t

$Y_{\{ij\}}^t$ Número máximo de ligações físicas do tipo t entre os nós i e j

$C_{\{ij\}}^t$ Custo associado à utilização de uma ligação do tipo t entre os nós i e j

S Conjunto de serviços s suportados pela rede

K_s Conjunto de VPCs k da rede lógica de suporte ao serviço s; cada VPC k é caracterizado por:

- P_k^s Conjunto dos trajectos candidatos p
- δ_{ij}^{psk} Constante binária: 1 indica que o trajecto candidato p passa em (i,j) da origem para o destino
- b_k^s Largura de banda no sentido da origem para o destino
- $\underline{b}_k^s$ Largura de banda no sentido do destino para a origem

$y_{\{ij\}}^t$ Variável inteira que determina o número de ligações físicas do tipo t entre os nós i e j

x_p^{sk} Variável binária: 1 significa que o trajecto candidato p é o trajecto solução do VPC k de suporte ao serviço s

Minimizar $$\sum_{\{i,j\}\in A}\sum_{t\in T} C_{\{ij\}}^t \cdot y_{\{ij\}}^t \qquad (2a)$$

Sujeito a:

$$\sum_{p\in P_k^s} x_p^{sk} = 1 \qquad , s \in S\ ,\ k \in K_s \qquad (2b)$$

$$\sum_{s\in S}\sum_{k\in K_s}\left(b_k^s\sum_{p\in P_k^s}\left(\delta_{ij}^{psk}\cdot x_p^{sk}\right)+\underline{b}_k^s\sum_{p\in P_k^s}\left(\delta_{ji}^{psk}\cdot x_p^{sk}\right)\right)\le\sum_{t\in T}\alpha_t\cdot y_{\{ij\}}^t \qquad ,(i,j)\in A \tag{2c}$$

$$y_{\{ij\}}^t \le Y_{\{ij\}}^t \qquad ,\{i,j\}\in A\,,\,t\in T \tag{2d}$$

$x_p^{sk}\in\{0,1\}$ $\quad y_{\{ij\}}^t$ inteiro não negativo

As restrições (2b) garantem que apenas um dos trajectos candidatos do VPC *k* é escolhido na solução final. As restrições (2c) garantem que a capacidade instalada em cada aresta é suficiente para suportar todos os VPCs que a atravessam nos seus dois sentidos. Finalmente as restrições (2d) garantem que o número máximo de ligações físicas não é ultrapassado na solução encontrada.

3.3 Modelo de Dimensionamento Multi-Horário

No modelo multi-horário, cada VPC é reconfigurado no início de cada período horário. No caso da modelização explícita, as variáveis binárias de trajecto associadas a cada VPC contemplam um índice adicional que identifica a que período horário se referem. No caso da modelização por trajectos candidatos, o mesmo acontece com as variáveis binárias de selecção de trajecto. Em ambos os casos, as capacidades dos VPCs (b_k^{sh} e $\underline{b}_k^{sh}$) têm valores diferentes em cada período e, por conseguinte, são previamente calculadas com base no tráfego esperado para cada um dos respectivos períodos horários.

Nos serviços de consulta, os modelos propostos levam em conta que os VPCs de um nó de tráfego devem terminar todos no mesmo nó servidor nos vários períodos horários (embora possam ter trajectos diferentes entre diferentes períodos). Esta é uma imposição operacional pois considera-se não ser aceitável que um cliente tenha de saber um endereço do servidor diferente em cada período horário definido. A reconfiguração dinâmica da rede deve ser transparente para os utilizadores.

Nas sub-secções seguintes propõem-se dois modelos de Programação Inteira cuja resolução permite dimensionar a rede física de menor custo e simultaneamente determinar os trajectos de todos os VPCs para todos os períodos horários na rede física resultante. O primeiro modelo proposto baseia-se na modelização explícita de VPCs e o segundo modelo proposto baseia-se na modelização de VPCs por trajectos candidatos.

3.3.1 Formulação por Modelização Explícita de VPCs

N Conjunto de nós da rede

A Conjunto de arestas $\{i,j\}$ da rede: (i,j) representa a aresta $\{i,j\}$ no sentido de i para j e (j,i) representa o sentido inverso

T Conjunto dos tipos t de ligações físicas possíveis de implementar entre os nós de rede

α_t Largura de banda de uma ligação física do tipo t

$Y^t_{\{ij\}}$ Número máximo de ligações físicas do tipo t entre os nós i e j

$C^t_{\{ij\}}$ Custo associado à utilização de uma ligação do tipo t entre os nós i e j

S Conjunto de serviços s suportados pela rede

H Conjunto de períodos horários h

K_s Conjunto de VPCs k da rede lógica de suporte ao serviço s; cada VPC k é caracterizado por:

- o^s_k Nó origem
- D^s_k Conjunto de possíveis nós destinos
- b^{sh}_k Largura de banda no sentido da origem para o destino no período horário h
- $\underline{b}^{sh}_k$ Largura de banda no sentido do destino para a origem no período horário h

$y^t_{\{ij\}}$ Variável inteira que determina o número de ligações físicas do tipo t entre os nós i e j

x^{skh}_{ij} Variável binária de trajecto: 1 significa que o VPC k de suporte ao serviço s passa no arco (i,j) no período horário h

Minimizar
$$\sum_{\{i,j\}\in A}\sum_{t\in T} C^t_{\{ij\}} \cdot y^t_{\{ij\}} \qquad (3a)$$

Sujeito a:
$$\sum_{j\in N:\{i,j\}\in A}\left(x^{skh}_{ij} - x^{skh}_{ji}\right)=1 \qquad , s\in S, k\in K_s, h\in H, i = o^s_k \qquad (3b)$$

$$\sum_{i\in D^s_k}\sum_{j\in N:\{i,j\}\in A}\left(x^{skh}_{ji} - x^{skh}_{ij}\right)=1 \qquad , s\in S, k\in K_s, h\in H \qquad (3c)$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x^{skh}_{ij} - x^{skh}_{ji}\right)=0 \qquad , s\in S, k\in K_s, h\in H, i\in N\setminus\{o^s_k, D^s_k\} \qquad (3d)$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x^{skh_1}_{ji} - x^{skh_1}_{ij}\right)= \sum_{j\in N:\{i,j\}\in A}\left(x^{skh_2}_{ji} - x^{skh_2}_{ij}\right) \quad \text{para qualquer } \{h_1,h_2\}\in H$$

$$, s\in S, k\in K_s, i\in D^s_k \qquad (3e)$$

$$\sum_{s\in S}\sum_{k\in K_s}\left(b_k^{sh}\cdot x_{ij}^{skh}+\underline{b}_k^{sh}\cdot x_{ji}^{skh}\right)\le\sum_{t\in T}\alpha_t\cdot y_{\{ij\}}^t \qquad ,(i,j)\in A\,,\,h\in H \tag{3f}$$

$$y_{\{ij\}}^t\le Y_{\{ij\}}^t \qquad ,\{i,j\}\in A\,,\,t\in T \tag{3g}$$

$x_{ij}^{skh}\in\{0,1\}$ $y_{\{ij\}}^t$ inteiro não negativo

As restrições (3b), (3c) e (3d) são as restrições conhecidas para VPCs com destino indefinido. Note-se que pelas restrições (3c) apenas um dos nós pertencentes a D_k^s poderá ter, em cada período horário *h*, a soma dos arcos de saída menos a soma dos arcos de entrada a um. Ao introduzirmos as restrições (3e) em cada um dos nós pertencentes a D_k^s, garantimos que o nó escolhido como destino para cada VPC é igual em todos os períodos horários. As restrições (3e) aplicam-se apenas aos VPCs em que o conjunto D_k^s não é unitário. Para os todos os outros VPCs, estas restrições são garantidas pelas restrições (3c) e são, por isso, redundantes. As restrições (3f) garantem que a capacidade instalada em cada aresta é suficiente para suportar todos os VPCs que por lá passam em ambos os sentidos e para todos os períodos horários. Finalmente as restrições (3g) garantem que o número máximo de ligações físicas não é ultrapassado na solução encontrada.

3.3.2 Formulação por Trajectos Candidatos

N Conjunto de nós da rede

A Conjunto de arestas $\{i,j\}$ da rede: (i,j) representa a aresta $\{i,j\}$ no sentido de *i* para *j* e (j,i) representa o sentido inverso

T Conjunto dos tipos *t* de ligações físicas possíveis de implementar entre os nós de rede

α_t Largura de banda de uma ligação física do tipo *t*

$Y_{\{ij\}}^t$ Número máximo de ligações físicas do tipo *t* entre os nós *i* e *j*

$C_{\{ij\}}^t$ Custo associado à utilização de uma ligação do tipo *t* entre os nós *i* e *j*

S Conjunto de serviços *s* suportados pela rede

H Conjunto de períodos horários *h*

K_s Conjunto de VPCs *k* da rede lógica de suporte ao serviço *s*; cada VPC *k* é caracterizado por:

- P_k^s Conjunto dos trajectos candidatos *p*

δ_{ij}^{psk} Constante binária: 1 indica que o trajecto candidato *p* passa em (*i*,*j*) da origem para o destino

b_k^{sh} Largura de banda no sentido da origem para o destino no período horário *h*

$\underline{b}_k^{sh}$ Largura de banda no sentido do destino para a origem no período horário *h*

$y_{\{ij\}}^t$ Variável inteira que determina o número de ligações físicas do tipo *t* entre os nós *i* e *j*

x_p^{skh} Variável binária: 1 significa que o trajecto candidato *p* é o trajecto solução no período horário *h* do VPC *k* de suporte ao serviço *s*

Minimizar $$\sum_{\{i,j\}\in A}\sum_{t\in T} C_{\{ij\}}^t \cdot y_{\{ij\}}^t \qquad (4a)$$

Sujeito a:

$$\sum_{p\in P_k^s} x_p^{skh} = 1 \qquad , s \in S\,,\, k \in K_s\,,\, h \in H \quad (4b)$$

$$\sum_{j\in N:\{i,j\}\in A}\sum_{p\in P_k^s}\left(\delta_{ji}^{psk}\cdot x_p^{skh_1}\right) = \sum_{j\in N:\{i,j\}\in A}\sum_{p\in P_k^s}\left(\delta_{ji}^{psk}\cdot x_p^{skh_2}\right) \quad \text{para } \{h_1,h_2\} \in H$$

$$, s \in S\,,\, k \in K_s\,,\, i \in D_k^s \quad (4c)$$

$$\sum_{s\in S}\sum_{k\in K_s}\left(b_k^{sh}\sum_{p\in P_k^s}\left(\delta_{ij}^{psk}\cdot x_p^{skh}\right)+\underline{b}_k^{sh}\sum_{p\in P_k^s}\left(\delta_{ji}^{psk}\cdot x_p^{skh}\right)\right)\le\sum_{t\in T}\alpha_t\cdot y_{\{ij\}}^t \quad , (i,j)\in A\,,\, h\in H \quad (4d)$$

$$y_{\{ij\}}^t \le Y_{\{ij\}}^t \qquad , \{i,j\}\in A\,,\, t\in T \qquad (4e)$$

$x_p^{skh} \in \{0,1\}$ $y_{\{ij\}}^t$ inteiro não negativo

As restrições (4b) garantem que apenas um dos trajectos candidatos do VPC *k* é escolhido na solução final. As restrições (4c) garantem que cada VPC dos serviços de consulta termina no mesmo nó destino (pertencente a D_k^s) para todos os períodos horários. Estas restrições são válidas porque não existe nenhum trajecto candidato que para chegar a um nó pertencente a D_k^s passe por outro nó pertencente a D_k^s. Assim, para cada período horário *h*, o termo

$$\sum_{j\in N:\{i,j\}\in A}\sum_{p\in P_k^s}\left(\delta_{ji}^{psk}\cdot x_p^{skh}\right)$$

é igual a um apenas num dos nós $i \in D_k^s$ e zero em todos os outros. Ao introduzirmos as restrições (4c) em cada um dos nós pertencentes a D_k^s, garantimos que o nó escolhido como destino para cada VPC *k* é igual em todos os períodos horários. As restrições (4c) só se aplicam aos VPCs em que o conjunto D_k^s não é unitário. Para os outros VPCs, estas restrições são garantidas pelo facto de todos os trajectos candidatos terminaram no mesmo nó destino e são, por isso, redundantes. As restrições (4d) garantem que a capacidade instalada em cada aresta é suficiente para suportar todos os VPCs que por lá passam em ambos os sentidos e para todos os períodos horários. Finalmente as restrições (4e) garantem que o número máximo de ligações físicas não é ultrapassado na solução encontrada.

3.4 Resolução dos Modelos de Dimensionamento

Os modelos propostos são modelos de Programação Inteira. Por conseguinte, é teoricamente possível serem resolvidos com ferramentas genéricas de resolução de modelos de Programação Inteira que calculam a solução óptima, como por exemplo o algoritmo de *branch-and-bound*. No entanto, este tipo de algoritmos não resolve em tempo útil os modelos propostos para problemas de dimensão média. A alternativa é a investigação de algoritmos heurísticos que calculem boas soluções sem que necessariamente sejam as soluções óptimas. O algoritmo aqui proposto baseia-se numa ferramenta da Investigação Operacional conhecida por Relaxação Lagrangeana com Optimização por Sub-Gradiente. De seguida, são primeiro revistos os principais conceitos associadas a esta ferramenta e depois é proposto o algoritmo de resolução dos modelos apresentados.

3.4.1 Relaxação Lagrangeana com Optimização por Sub-Gradiente

A relaxação Lagrangeana foi desenvolvida no princípio dos anos 70 com o trabalho pioneiro de Held e Karp ([Held70] e [Held71]) aplicado ao problema clássico em Investigação Operacional do caixeiro viajante. É uma técnica importante nos dias actuais para a geração de limites inferiores na resolução de problemas de optimização combinatória e permite também servir de base ao cálculo de soluções possíveis para os problemas. Considere-se o seguinte problema genérico (em notação matricial) com variáveis *x* inteiras:

Problema *P*: minimizar cx

sujeito a: $Ax \geq b$

$Bx \geq d$

$x \in \{0,1\}$

Define-se a relaxação Lagrangeana do problema *P* referente ao conjunto de restrições $Ax \geq b$ introduzindo um vector de multiplicadores de Lagrange $\lambda \geq 0$ aplicado a estas restrições e introduzindo-as na função objectivo da seguinte forma:

Problema *RL*: minimizar $cx + \lambda(b - Ax)$

sujeito a: $Bx \geq d$

$x \in \{0,1\}$

A solução óptima do problema *RL* (relaxação Lagrangeana) é um limite inferior da solução óptima do problema *P* dado que

o valor de minimizar cx

sujeito a: $Ax \geq b$

$Bx \geq d$

$x \in \{0,1\}$

é maior ou igual do que o valor de minimizar $cx + \lambda(b - Ax)$

sujeito a: $Ax \geq b$

$Bx \geq d$

$x \in \{0,1\}$

(porque $\lambda \geq 0$ e $b - Ax \leq 0$ e portanto estamos a introduzir um termo menor que zero na função objectivo)

e é também maior ou igual do que o valor de minimizar $cx + \lambda(b - Ax)$

sujeito a: $Bx \geq d$

$x \in \{0,1\}$

(porque retirar um conjunto de restrições significa aumentar o conjunto de soluções admissíveis do problema pelo que o valor da função objectivo da solução ou se mantém ou é reduzido).

A formulação da relaxação Lagrangeana envolve dois aspectos importantes. O primeiro aspecto é o da selecção das restrições envolvidas na relaxação. Note-se que no caso anterior podemos definir outra relaxação Lagrangeana se envolvermos as restrições

B$x \geq$ d. Um dos critérios na selecções das restrições é o da obtenção de um problema de optimização mais simples de resolver do que o problema original *P*. O segundo aspecto importante é o da selecção dos valores dos multiplicadores de Lagrange λ. Dado que a solução da relaxação Lagrangeana determina um limite inferior do problema original, estamos interessados em escolher os multiplicadores de Lagrange que maximizam o valor da função objectivo da relaxação Lagrangeana. Este problema de maximização é designado por problema dual Lagrangeano.

Se o valor de uma solução do problema original for igual ao valor da solução óptima do seu problema dual Lagrangeano, então sabemos que encontrámos a solução óptima do problema original. No caso geral, isso não acontece e diz-se que existe um intervalo dual entre a solução do problema original e o limite inferior calculado pelo problema dual Lagrangeano. O intervalo dual é dado pela diferença relativa entre os dois valores e permite-nos aferir da qualidade de uma solução encontrada para o problema original.

A optimização por sub-gradiente é um procedimento iterativo que, a partir de um conjunto de multiplicadores de Lagrange, calcula novos conjuntos de uma forma sistemática, tentando maximizar o valor da relaxação Lagrangeana ([Beasley93], [Held74]). Considere-se as restrições relaxadas na sua notação não matricial

$$\sum_{j=1}^{n} a_{ij} x_j \geq b_i \qquad i = 1,2,\ldots,m$$

Considere-se P_{LS} um limite superior para o problema original *P* (dado pelo valor de uma solução admissível do problema original). Considere-se inicialmente $l = 0$ e considere um conjunto inicial de multiplicadores de Lagrange $\Lambda_0 = \{ \lambda_1^0, \lambda_2^0, \ldots, \lambda_m^0 \}$. Seja $RL(\Lambda_l)$ o valor da função objectivo da solução do problema *RL* para o conjunto de multiplicadores de Lagrange Λ_l e $X(\Lambda_l) = \{x_i(\Lambda_l),\ i = 1,2,\ldots,n\}$, os respectivos valores da solução. Considere-se ainda um escalar π com o valor inicial de 2. A optimização por sub-gradiente é dada pelo seguinte processo iterativo:

(i) Calcular $X(\Lambda_l)$ e o limite inferior de *P* dado por $P_{LI} = RL(\Lambda_l)$.

(ii) Definir sub-gradientes G_i da solução calculada $X(\Lambda_l)$ da seguinte forma:

$$G_i = b_i - \sum_{j=1}^{n} a_{ij} \cdot x_j(\Lambda_l)\ i = 1,2,\ldots,m$$

(iii) Definir um passo *T* da seguinte forma:

$$T = \pi \cdot \frac{P_{LS} - P_{LI}}{\sum_{i=1}^{m} (G_i)^2}$$

(iv) Fazer $l = l + 1$, recalcular os multiplicadores de Lagrange Λ_l da seguinte forma:

$$\lambda_i^l = \max(0, \lambda_i^{l-1} + T \cdot G_i) \qquad i = 1,2,\ldots,m$$

e saltar para o passo (i).

Existem duas alternativas para terminar o processo. A primeira é executar o processo até um número pré-determinado de iterações. A segunda é fazer o valor π ir diminuindo sempre que o valor de P_{LI} não melhora ao fim de um número pré-determinado de iterações e parar o processo quando π for suficientemente pequeno. Um critério típico usado em muitos casos na literatura é reduzir o valor de π para metade sempre que P_{LI} não melhora ao fim de 40 iterações.

3.4.2 Algoritmo Implementado

Considere-se o problema multi-horário formulado por modelização explícita de VPCs apresentado na secção 3.3.1 (o problema uni-horário formulado apresentado na secção 3.2.1 é um caso particular do problema multi-horário em que o número de períodos horários é um, pelo que a descrição que se segue também se aplica a este caso). Este modelo é complexo de resolver por causa das restrições (3f) que relacionam entre si as variáveis *x* e *y*. Se estas restrições forem relaxadas, o problema resultante é de resolução fácil. Assim, partindo do modelo original, define-se um novo problema de optimização aplicando a relaxação Lagrangeana às restrições (3f). Este novo problema tem a seguinte formulação:

$RL(\Lambda)$: $P1(\Lambda) + P2(\Lambda)$ (5)

em que

$P1(\Lambda)$: Minimizar $\sum_{\{i,j\} \in A} \sum_{t \in T} \left(C_{\{ij\}}^t - \alpha_t \sum_{h \in H} \left(\lambda_{ij}^h + \lambda_{ji}^h \right) \right) y_{\{ij\}}^t$ (6a)

Sujeito a:

$y_{\{ij\}}^t < Y_{\{ij\}}^t \qquad , \{i,j\} \in A \, , t \in T$ (6b)

$y_{\{ij\}}^t$ inteiro não negativo

P2(Λ): Minimizar $\sum_{s\in S}\sum_{k\in K_s}\sum_{h\in H}\sum_{\{i,j\}\in A}\left(\lambda_{ij}^{h}\cdot b_k^{sh}+\lambda_{ji}^{h}\cdot \underline{b}_k^{sh}\right)x_{ij}^{skh}$ (7a)

Sujeito a:

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{ij}^{skh}-x_{ji}^{skh}\right)=1 \qquad ,s\in S\,,k\in K_s\,,h\in H\,,i=o_k^s \tag{7b}$$

$$\sum_{i\in D_k^s}\sum_{j\in N:\{i,j\}\in A}\left(x_{ji}^{skh}-x_{ij}^{skh}\right)=1 \qquad ,s\in S\,,k\in K_s\,,h\in H \tag{7c}$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{ij}^{skh}-x_{ji}^{skh}\right)=0 \qquad ,s\in S\,,k\in K_s\,,h\in H\,,i\in N\setminus\{o_k^s,D_k^s\} \tag{7d}$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{ji}^{skh_1}-x_{ij}^{skh_1}\right)=\sum_{j\in N:\{i,j\}\in A}\left(x_{ji}^{skh_2}-x_{ij}^{skh_2}\right) \quad \text{para qualquer } \{h_1,h_2\}\in H$$

$$,s\in S\,,k\in K_s\,,i\in D_k^s \tag{7e}$$

$$x_{ij}^{skh}\in\{0,1\}$$

Para enunciar o problema *RL*(Λ), foi introduzido um conjunto multiplicadores de Lagrange $\Lambda=\{\lambda_{ij}^{h},\ (i,j)\in A,\ h\in H\}$, um por cada restrição relaxada. Para qualquer conjunto Λ de multiplicadores de Lagrange, o valor da função objectivo da solução de *RL*(Λ) é um limite inferior da solução do problema original. A solução de *RL*(Λ) é a soma das soluções dos sub-problemas *P1*(Λ) e *P2*(Λ) que são facilmente calculáveis. A solução de *P1*(Λ) é dada por:

$$y(\Lambda)=\left\{\hat{y}_{\{ij\}}^{t}\right\} \qquad ,\{i,j\}\in A\,,t\in T$$

em que

$$\hat{y}_{\{ij\}}^{t}=\begin{cases}0 & ,C_{ij}^{t}\geq\alpha_t\cdot\sum_{h\in H}\left(\lambda_{ij}^{h}+\lambda_{ji}^{h}\right)\\ Y_{\{ij\}}^{t} & ,C_{ij}^{t}<\alpha_t\cdot\sum_{h\in H}\left(\lambda_{ij}^{h}+\lambda_{ji}^{h}\right)\end{cases} \tag{8}$$

O sub-problema *P2*(Λ) é ainda decomponível numa soma de problemas elementares de trajecto mais curto. Para cada VPC $k\in K_s$ de cada serviço $s\in S$, em cada período horário $h\in H$, considera-se o grafo original (*N*,*A*) em que a distância de cada arco é dada por:

$$d_{ij}^{skh}=\lambda_{ij}^{h}\cdot b_k^{sh}+\lambda_{ji}^{h}\cdot\underline{b}_k^{sh} \tag{9}$$

e a solução de *P2*(Λ) é dada pela aplicação de um algoritmo de trajecto mais curto. Considere-se esta solução definida por:

$$x(\Lambda) = \left\{\hat{x}_{ij}^{skh}\right\} \qquad , s \in S\ , k \in K_s\ , h \in H\ , (i,j) \in A \tag{10}$$

Existem algoritmos de complexidade polinomial para resolver o sub-problema *P2*(Λ). O algoritmo de Dijkstra, usado na presente implementação, calcula o trajecto mais curto de um nó origem para todos os outros nós do grafo. Para os VPCs com destino indefinido, o algoritmo de Dijkstra permite também escolher simultaneamente, de entre os nós de D_k^s, o nó destino que resulta no trajecto mais curto. Basta para isso fazer terminar o algoritmo quando os trajectos mais curtos para todos os nós de D_k^s tiverem sido processados. Quando o número de períodos horários é maior que um, a determinação dos trajectos dos VPCs com destino indefinido é um pouco mais complexa porque é preciso garantir que os trajectos referentes a uma mesma origem têm o mesmo nó de D_k^s como destino para os diferentes períodos horários. A solução implementada começa por determinar os trajectos mais curtos para todos os possíveis destinos em todos os períodos horários. De seguida, para cada possível destino são somadas as distâncias dos trajectos de todos os períodos horários. Finalmente, escolhe o nó destino (e os respectivos trajectos) cuja soma das distâncias dos trajectos para todos os períodos horários seja mínima.

Os valores dados por (8) e (10) definem a solução do problema *RL*(Λ) para um conjunto particular de multiplicadores de Lagrange Λ e a equação (5) define o valor da função objectivo da solução. De seguida, determina-se uma solução admissível do problema original. Este passo é efectuado considerando $x_{ij}^{skh} = \hat{x}_{ij}^{skh}$ e usando as restrições (3f) para calcular as variáveis $y_{\{ij\}}^t$ que minimizam (3a). Esta operação é constituída por um conjunto de problemas de optimização, um por cada aresta $\{i,j\}$, com a seguinte formulação:

Minimizar $$\sum_{t \in T} C_{\{ij\}}^t \cdot y_{\{ij\}}^t \tag{11a}$$

Sujeito a:

$$\sum_{t \in T} \alpha_t \cdot y_{\{ij\}}^t \geq M_{\{ij\}} \tag{11b}$$

$$y_{\{ij\}}^t \leq Y_{\{ij\}}^t \qquad , t \in T \tag{11c}$$

em que $M_{\{ij\}}$ é o valor da capacidade total que passa pela aresta $\{i,j\}$ no sentido e período horário mais exigentes, ou seja:

$$M_{\{ij\}} = \max_{h \in H}\left(M_{\{ij\}}^{h}\right) \qquad \text{em que:}$$

$$M_{\{ij\}}^{h} = \max\left(\sum_{s \in S} \sum_{k \in K_s} \left(b_k^{sh} \cdot x_{ij}^{skh} + \underline{b}_k^{sh} \cdot x_{ji}^{skh} \right) , \sum_{s \in S} \sum_{k \in K_s} \left(b_k^{sh} \cdot x_{ji}^{skh} + \underline{b}_k^{sh} \cdot x_{ij}^{skh} \right) \right) \qquad (11d)$$

Assim, dado um conjunto Λ de multiplicadores de Lagrange, calculamos um limite inferior, P_{LI}, através da solução do problema $RL(\Lambda)$, e uma solução possível do problema original que é também um limite superior, P_{LS}, da solução do problema original. A diferença destes dois valores dá uma medida da qualidade da solução encontrada.

Para calcular diferentes conjuntos de multiplicadores de Lagrange, usamos a optimização por sub-gradiente. Como foi atrás explicado, esta técnica é um processo iterativo que, para um determinado conjunto Λ_n, calcula um novo conjunto Λ_{n+1} na tentativa de maximizar o valor da função objectivo de problema *RL*. Esta operação é feita da seguinte forma. Primeiro, um valor de sub-gradiente é calculado para cada restrição relaxada:

$$G_{ij}^{h} = \sum_{s \in S} \sum_{k \in K_s} \left(b_k^{sh} \cdot \hat{x}_{ij}^{skh} + \underline{b}_k^{sh} \cdot \hat{x}_{ji}^{skh} \right) - \sum_{t \in T} \alpha_t \cdot \hat{y}_{\{ij\}}^{t} \quad , (i,j) \in A \, , \, t \in T \qquad (12)$$

Depois, é definido o passo *T*:

$$T = \pi \cdot \frac{P_{LS} - P_{LI}}{\sum_{h \in H} \sum_{(i,j) \in A} \left(G_{ij}^{h} \right)^2} \qquad (13)$$

em que P_{LI} é o valor do limite inferior actual (da solução actual de $RL(\Lambda)$) e P_{LS} o limite superior (valor da melhor solução do problema original). π é um parâmetro da relaxação ao qual é atribuído inicialmente o valor 2. Finalmente, um novo conjunto de multiplicadores de Lagrange é calculado com base no conjunto anterior da seguinte maneira:

$$\lambda_{ij}^{h} = \max\left(0, \lambda_{ij}^{h} + T \cdot G_{ij}^{h}\right) \qquad (14)$$

Na implementação desenvolvida, parametrizamos o algoritmo para executar um total de 1000 iterações e o valor de π é dividido por dois sempre que o valor de P_{LI} não melhora em 40 iterações consecutivas. Durante o processo iterativo, são calculadas 1000 soluções admissíveis do problema original, das quais a de menor custo constitui o resultado final do

algoritmo. De forma equivalente, o limite inferior para o problema original é dado pelo maior de todos os 1000 valores de P_{LI}.

Para que a descrição do algoritmo implementado esteja completa, falta apenas descrever como foi implementada a resolução dos problemas de optimização, definidos em (11), referentes ao cálculo de uma solução admissível no problema original. Cada um destes problemas é facilmente resolvido por procura exaustiva dado que nos casos práticos de interesse o número de variáveis é muito pequeno. Os casos reais consideram tipicamente um número máximo de tipos de interfaces de 4 ($|T| = 4$). No entanto, a eficiência computacional do algoritmo é melhorada se previamente forem calculadas para cada aresta, todas as combinações possíveis de valores das variáveis $y^t_{\{ij\}}$ e ordenadas pela sua capacidade total. Destas, são ainda eliminadas as combinações para as quais haja outras com maior capacidade cujo custo não seja maior. Deste modo, a resolução de cada um dos problemas (11) é implementada por uma operação de consulta a uma tabela ordenada, previamente calculada, em que é escolhida a primeira combinação encontrada cuja capacidade é maior ou igual que $M_{\{ij\}}$. Por exemplo, considere-se que para uma determinada aresta $\{i,j\}$ existem 2 tipos de interfaces I1 e I2 caracterizadas por:

$\alpha_1 = 10.0$ $\quad C^1_{\{ij\}} = 225.0$ $\quad Y^1_{\{ij\}} = 4$

$\alpha_2 = 40.0$ $\quad C^2_{\{ij\}} = 500.0$ $\quad Y^2_{\{ij\}} = 4$

Embora o número máximo de interfaces do tipo I1 seja 4, verifica-se que três interfaces deste tipo custam 675.0 e têm uma capacidade de 30.0 enquanto que uma única interface do tipo I2 custa menos (custo = 500.0) e tem mais capacidade (capacidade = 40.0). Assim, a combinação de 3 interfaces do tipo I1 nunca será uma solução de menor custo para esta aresta e não precisa de pertencer à lista de combinações de interfaces. A tabela ordenada resultante para a aresta $\{i,j\}$ é:

Entrada	Capacidade	Custo	Combinação
1	10.0	225.0	$1 \times$ I1
2	20.0	450.0	$2 \times$ I1
3	40.0	500.0	$1 \times$ I2
4	50.0	725.0	$1 \times$ I1 , $1 \times$ I2
5	60.0	950.0	$2 \times$ I1 , $1 \times$ I2

6	80.0	1000.0	$2 \times I2$
…	…	…	…

Esta tabela previamente calculada para a aresta $\{i,j\}$ é consultada em todas as iterações do algoritmo. Se por exemplo numa determinada iteração $M_{\{ij\}}$ = 53.45, o algoritmo procura sequencialmente na coluna da capacidade o primeiro valor maior que 53.45, e encontra na entrada 4 o custo de 725.0 para as interfaces a instalar nesta aresta. Esta operação é feita para todas as arestas e o custo da solução é dado pela soma dos custos de todas as arestas. Esta estratégia diminui o tempo total de cálculo à custa de uma maior exigência em termos de memória computacional o que nos sistemas computacionais actuais não é crítico porque o custo de memória é substancialmente inferior ao custo de capacidade de processamento.

O algoritmo descrito aplica-se aos modelos baseados na modelização explícita de VPCs. No entanto, este algoritmo é facilmente adaptado aos modelos baseados em trajectos candidados para os VPCs. O algoritmo mantém-se conforme foi descrito e é apenas alterado na resolução do problema *P2*(Λ). Como os trajectos para cada VPC $k \in K_s$ em cada serviço $s \in S$ são previamente determinados, a resolução do problema *P2*(Λ) é dada pela escolha do trajecto que, de entre os trajectos candidatos, minimiza a distância da origem para o destino em que as distâncias de cada arco são dadas por (9). Tal como no caso anterior, a escolha dos trajectos dos VPCs com destino indefinido é um pouco mais complexa para o modelo multi-horário porque é preciso garantir que os trajectos escolhidos para uma mesma origem têm o mesmo nó destino nos diferentes períodos horários. A solução implementada é, no entanto, semelhante ao caso anterior. São escolhidos os trajectos (um por período horário) para o nó tal que a soma das suas distâncias seja a menor.

3.5 Resultados Computacionais

Com o objectivo de estudar o desempenho do algoritmo heurístico proposto, o algoritmo proposto foi implementado em linguagem C++ para uma plataforma PC com sistema operativo Windows da Microsoft. Para uma melhor compreensão à análise dos resultados computacionais obtidos nos diferentes aspectos abordados, esta secção está organizada da seguinte forma. A secção 3.5.1 descreve os problemas considerados para

dimensionamento usando o algoritmo proposto. Esta secção aborda a forma como foram determinadas as topologia das redes, os custos de transmissão e de comutação e os cenários de tráfego. A secção 3.5.2 faz uma análise comparativa da resolução dos problemas de dimensionamento entre o algoritmo baseado na modelização explícita de VPCs e o algoritmo baseado na modelização dos VPCs por trajectos candidatos. A secção 3.5.3 compara os resultados de dimensionamento entre os modelos uni-horário e multi-horário. Na secção 3.5.4, é feita uma avaliação do desempenho do algoritmo proposto em termos da sua eficiência computacional (tempos de cálculo) e eficácia (qualidade das soluções obtidas). Na secção 3.5.5, com base na proposta de um método alternativo para a determinação de limites inferiores, é feita uma avaliação da qualidade dos limites inferiores determinados pelo algoritmo proposto. Finalmente, a secção 3.5.6 apresenta as principais conclusões obtidas com base noutros resultados computacionais efectuados mas não apresentados nesta tese.

3.5.1 Definição dos Problemas Considerados

Com o objectivo de estudar o desempenho do algoritmo heurístico proposto, consideraram-se diferentes redes com um número de nós entre 10 e 50 (intervalos de 5 em 5 nós). Para cada um destes valores, geraram-se 2 redes: (i) uma rede esparsa com um número de arestas dado pelo menor inteiro que seja maior ou igual a 1.5 vezes o número de nós e (ii) uma rede densa com um número de arestas dado pelo menor inteiro que seja maior ou igual a 2.5 vezes o número de nós. A designação de cada uma das redes e as suas grandezas topológicas são apresentadas na Tabela 3-1.

As redes foram geradas do modo que a seguir se descreve. Considerou-se um espaço geográfico rectangular de 240 Km por 180 Km para todas as redes. Geraram-se aleatoriamente 10 localizações para os nós que servem de base às redes 10S e 10D e geraram-se aleatoriamente arestas para a rede 10S garantindo no entanto que cada nó é extremo de pelo menos duas arestas. Determinada a rede 10S, gerou-se a rede 10D acrescentando aleatoriamente as restantes arestas. Às arestas geradas é atribuído um comprimento igual à distância Euclidiana entre os seus nós extremo. Com base na rede 10S, gerou-se a rede 15S acrescentando aleatoriamente os nós e as arestas necessárias. A partir daqui, as restantes redes foram geradas segundo os mesmos critérios das anteriores. A topologia das redes geradas é apresentada no Anexo B desta tese.

Designação Da Rede	Número De Nós	Número de Arestas
10D	10	25
10S	10	15
15D	15	38
15S	15	23
20D	20	50
20S	29	30
25D	25	63
25S	25	38
30D	30	75
30S	30	45
35D	35	88
35S	35	53
40D	40	100
40S	40	60
45D	45	113
45S	45	68
50D	50	125
50S	50	75

Tabela 3-1 : Informação topológica das redes estudadas

Para todas as redes, considerou-se a possibilidade de usar ligações físicas de 3 tipos (as capacidades apresentadas são as capacidades de transmissão de informação das tecnologias consideradas):

- Tipo 1: ligações PDH E3 (α_1 = 3.392×10^7 bps)
- Tipo 2: ligações SDH STM-1 (α_2 = 1.4976×10^8 bps)
- Tipo 3: ligações SDH STM-4 (α_3 = 5.9904×10^8 bps)

Os custos $C^t_{\{ij\}}$ foram calculados atribuíndo um custo de comutação de 10 (ligações PDH E3), 35 (ligações SDH STM-1) e 130 (ligações SDH STM-1) para todos os nós e considerando ligações dedicadas com um custo de transmissão de 1 por kilómetro. O valor absoluto dos custos atribuídos não tem significado real (daí não serem atribuídas unidades aos valores) mas o valor relativo entre eles foi escolhido segundo os critérios que se descrevem de seguida.

Note-se que a capacidade suportada por 4 ligações PDH E3 é sensivelmente igual à capacidade suportada por 1 ligação SDH STM-1 (a mesma relação existe entre as ligações SDH STM-1 e SDH STM-4). Como um comutador com 4 interfaces PDH E3 é mais

complexo do que um comutador com 1 interface SDH STM-1 (em termos de hardware e de gestão do processamento da informação), então os custos de comutação de 4 interfaces PDH E3 deverão ser maiores do que os custos de comutação de uma interface SDH STM-1. Assim, escolhido arbitrariamente o custo de comutação de 10 para as interfaces do tipo 1, foi escolhido o valor de 35 para as interfaces do tipo 2 (menor de que 40 para 4 interfaces do tipo 1) e de 130 para as interfaces do tipo 3 (menor de que 140 para 4 interfaces do tipo 2).

Em relação aos custos de transmissão, relembre-se que (secção 2.2.3) cada valor $C^t_{\{ij\}}$ é dado por 2 vezes o custo de comutação do tipo *t* mais o custo de transmissão associado à distância da aresta $\{i,j\}$. Fazendo os cálculos adequados, verifica-se que um custo de transmissão de 1 por kilómetro conduz ao seguinte cenário:

- Para ligações menores ou iguais a 5 kilómetros, 3 ligações de tipo 1 têm menor custo que uma ligação de tipo 2; para ligações entre 5 e 30 kilómetros, 2 ligações de tipo 1 têm menor custo que uma ligação de tipo 2; para ligações maiores que 30 kilómetros, apenas 1 ligação do tipo 1 tem menor custo que uma ligação de tipo 2.
- Para ligações menores ou iguais a 25 kilómetros, 3 ligações de tipo 2 têm menor custo que uma ligação de tipo 3; para ligações entre 25 e 120 kilómetros, 2 ligações de tipo 2 têm menor custo que uma ligação de tipo 3; para ligações maiores que 120 kilómetros, apenas 1 ligação do tipo 2 tem menor custo que uma ligação de tipo 3.

Assim, o valor do custo de transmissão escolhido (1 por kilómetro) diferencia o número máximo de tipos de ligações intermédias entre ligações de curta distância e longa distância de uma forma considerada realista. Um valor muito menor faria com que o número máximo de ligações do tipo 1 e 2 fosse três para qualquer distância e um valor muito maior faria com que o número máximo de ligações do tipo 1 e 2 fosse um para qualquer distância.

Para o dimensionamento de cada uma das redes, consideramos 10 cenários de tráfego multi-horário com 3 períodos horários e 3 hipotéticos serviços conversacionais: serviço telefónico a 64 Kbps por ligação, serviço de video-telefonia a 256 Kbps por ligação e serviço de transmissão de dados a 2 Mbps por ligação. A determinação de cada cenário de tráfego foi feita da seguinte forma:

(i) determinação dos nós de tráfego de cada serviço escolhendo aleatoriamente 80% dos nós da rede;

(ii) geração aleatória dos valores de tráfego para cada par de nós de tráfego de todos os serviços e todos os períodos horários com uma função densidade de probabilidade uniforme entre 0 e 90 erlangs (telefonia), 0 e 40 erlangs (video-telefonia) e 0 e 12 erlangs (dados).

De cada cenário de tráfego multi-horário, gerou-se um cenário uni-horário em que para cada par de nós de tráfego de cada serviço, foi considerado o maior dos valores de tráfego dos três períodos horários. O objectivo é avaliar, para os cenários de tráfego gerados, quais são os ganhos que se obtêm, em termos de custo da rede dimensionada, se a rede permitir reconfiguração dinâmica dos VPCs.

Assim, consideramos um total de 360 problemas resultante de 18 redes com 10 problemas uni-horário e 10 problemas multi-horário para cada rede. Todos os problemas foram resolvidos pelo algoritmo descrito na secção anterior com modelização explícita de VPCs. Os valores do custo da melhor solução encontrada, do limite inferior para o custo da solução óptima e do tempo de cálculo para os 360 problemas são apresentados no Anexo B desta tese. Usamos também o algoritmo descrito na secção anterior baseado em trajectos candidatos para resolver os problemas das redes com 10, 15, 30 e 45 nós (um total de 8 redes e 20 problemas em cada rede). Resolvemos estes problemas considerando duas alternativas: 5 trajectos candidatos e 20 trajectos candidatos. Os trajectos candidatos foram determinados pelo algoritmo proposto em [Martins97]. Tal como no caso anterior, os valores do custo da melhor solução encontrada, do limite inferior para o custo da solução óptima e do tempo de cálculo para os 160 problemas são apresentados no Anexo B.

3.5.2 Modelização Explícita vs. Modelização por Trajectos Condidatos

Um dos objectivos dos problemas propostos é comparar o desempenho da modelização explícita de VPCs com a modelização por trajectos candidatos. Esta comparação deverá ter em conta não só o custo da melhor solução encontrada por cada um dos dois modelos mas também o tempo de execução dos respectivos algoritmos. Os resultados que a seguir se apresentam são relativos aos valores dados pela modelização explícita de VPCs. Por exemplo, um resultado de 10% para um determinado problema significa que esse valor (quer seja custo da solução encontrada, quer seja tempo de cálculo) é dez por cento maior no modelo baseado em trajectos candidatos do que no modelo

baseado em modelização explícita de VPC (um valor de percentagem negativo significa um valor menor).

Consideremos em primeiro lugar o caso de 5 trajectos candidatos. A Tabela 3-2 apresenta os valores mínimo, médio e máximo da comparação do custo das melhores soluções dadas pelos dois modelos nos 10 problemas de cada caso. A coluna *Melhores* especifica de entre os 10 problemas, em quantos é que a modelização explícita de VPCs deu um custo menor. A Tabela 3-3 apresenta os resultados comparativos para os tempos de cálculo de cada modelo.

Os resultados apresentados mostram que o algoritmo com modelização explícita de VPCs é superior ao algoritmo baseado em trajectos candidatos quando se consideram 5 trajectos candidatos para cada VPC. As melhores soluções do algoritmo baseado em trajectos candidatos têm um custo médio superior ao algoritmo com modelização explícita de VPCs para todas as redes e nos dois casos horários e em apenas 13 dos 160 problemas conseguiu encontrar uma solução de custo menor. Pelos resultados apresentados na Tabela 3-2, a modelização por trajectos candidatos calcula soluções que são em média 8.76% piores que a modelização explícita de VPCs.

Em relação aos tempos de cálculo, verifica-se que a modelização explícita de VPCs é mais eficiente para as redes de 30 e 45 nós (aquelas em que os tempos de cálculo são significativos). Para as redes pequenas (10 e 15 nós no caso uni-horário e 10 nós no caso multi-horário) os tempos de cálculo são piores mas estes casos não são críticos porque correspondem a valores absolutos de tempo de cálculo nunca superiores a 12 segundos.

Rede	Mínimo	Médio	Máximo	Melhores	Rede	Mínimo	Médio	Máximo	Melhores
10S	-2,66%	5,36%	12,23%	8	10D	9,10%	16,43%	23,94%	10
15S	-8,76%	0,17%	3,73%	7	15D	10,15%	13,74%	19,93%	10
30S	-0,97%	0,26%	1,53%	6	30D	3,25%	4,79%	6,79%	10
45S	1,58%	2,71%	3,31%	10	45D	8,21%	9,10%	10,37%	10

Uni-Horário

Rede	Mínimo	Médio	Máximo	Melhores	Rede	Mínimo	Médio	Máximo	Melhores
10S	2,85%	10,60%	20,50%	10	10D	5,28%	34,39%	66,27%	10
15S	-3,14%	4,43%	12,26%	8	15D	7,57%	23,50%	37,67%	10
30S	-0,86%	1,01%	2,54%	9	30D	-0,11%	1,47%	4,07%	9
45S	2,35%	3,58%	4,56%	10	45D	5,83%	8,58%	10,73%	10

Multi-Horário

Tabela 3-2 : Comparação dos valores das melhores soluções entre a modelização explícita e a modelização com 5 trajectos candidatos

Rede	Mínimo	Médio	Máximo	Melhores	Rede	Mínimo	Médio	Máximo	Melhores
10S	-26,15%	-22,98%	-21,89%	0	10D	-14,50%	-4,30%	-0,15%	0
15S	-19,47%	-11,50%	-9,09%	0	15D	11,63%	13,08%	14,02%	10
30S	12,52%	13,25%	14,07%	10	30D	41,75%	42,72%	44,59%	10
45S	15,30%	19,51%	22,88%	10	45D	39,95%	44,15%	63,61%	10

Uni-Horário

Rede	Mínimo	Médio	Máximo	Melhores	Rede	Mínimo	Médio	Máximo	Melhores
10S	-37,67%	-36,47%	-35,05%	0	10D	-16,99%	-15,56%	-14,32%	0
15S	-20,43%	-18,74%	-16,24%	0	15D	1,04%	2,17%	4,35%	10
30S	-2,80%	2,17%	6,21%	9	30D	28,70%	30,38%	31,99%	10
45S	11,34%	12,61%	13,62%	10	45D	31,02%	33,90%	38,19%	10

Multi-Horário

Tabela 3-3 : Comparação dos tempos de cálculo entre a modelização explícita e a modelização com 5 trajectos candidatos

Consideremos agora a comparação dos dois modelos em que se consideraram 20 trajectos candidatos. A Tabela 3-4 e a Tabela 3-5 apresentam os resultados comparativos para os custos das soluções encontradas e os tempos de cálculo. Estes resultados mostram o efeito de se considerar um maior número de trajectos candidatos: os valores das soluções encontradas por este método melhoram mas os tempos de cálculo aumentam significativamente. O que se pretende avaliar é se estes dois efeitos fazem com que este método seja comparativamente melhor do que a modelização explícita de VPCs. Em termos globais, a modelização explícita de VPCs é melhor também neste caso porque a modelização baseada em trajectos candidatos exibe tempos de cálculo que são em média bastante superiores e os custos das soluções encontradas continuam a ser em média piores (neste caso, 3.87%) que os custos das soluções dadas pela modelização explícita de VPCs.

Rede	Mínimo	Médio	Máximo	Melhores	Rede	Mínimo	Médio	Máximo	Melhores
10S	-9,96%	-0,92%	10,85%	4	10D	0,68%	8,61%	15,53%	10
15S	-2,54%	1,21%	5,09%	7	15D	-1,39%	2,82%	7,32%	8
30S	-0,94%	0,08%	1,42%	6	30D	-0,04%	1,48%	2,96%	9
45S	-0,73%	-0,21%	0,29%	3	45D	3,20%	4,08%	5,18%	10

Uni-Horário

Rede	Mínimo	Médio	Máximo	Melhores	Rede	Mínimo	Médio	Máximo	Melhores
10S	2,85%	10,74%	16,33%	10	10D	3,51%	20,08%	68,28%	10
15S	-8,31%	-0,25%	9,03%	5	15D	-6,97%	8,06%	21,50%	7
30S	-0,93%	0,20%	1,05%	7	30D	-0,93%	0,92%	2,79%	8
45S	-0,64%	0,44%	3,38%	7	45D	2,17%	4,62%	5,92%	10

Multi-Horário

Tabela 3-4 : Comparação dos valores das melhores soluções entre a modelização explícita e a modelização com 20 trajectos candidatos

Rede	Mínimo	Médio	Máximo	Melhores	Rede	Mínimo	Médio	Máximo	Melhores
10S	107,65%	115,92%	125,25%	10	10D	157,18%	184,10%	194,12%	10
15S	165,82%	180,65%	190,85%	10	15D	245,74%	257,48%	264,33%	10
30S	259,18%	261,95%	264,91%	10	30D	330,80%	334,50%	337,52%	10
45S	249,44%	260,26%	268,02%	10	45D	320,89%	324,28%	326,64%	10

Uni-Horário

Rede	Mínimo	Médio	Máximo	Melhores	Rede	Mínimo	Médio	Máximo	Melhores
10S	96,98%	99,12%	101,27%	10	10D	171,30%	182,24%	242,13%	10
15S	153,80%	159,86%	167,38%	10	15D	222,73%	229,72%	246,20%	10
30S	215,91%	232,37%	237,99%	10	30D	302,60%	305,34%	308,02%	10
45S	239,56%	242,67%	246,17%	10	45D	293,47%	304,47%	316,79%	10

Multi-Horário

Tabela 3-5 : Comparação dos tempos de cálculo entre a modelização explícita e a modelização com 20 trajectos candidatos

No entanto, existe um número significativo de problemas (39 em 160) em que o modelo baseado em trajectos candidatos deu uma solução melhor e a maioria destes casos acontece nas redes esparsas. Isto deve-se ao facto que 20 trajectos candidatos contemplam uma percentagem significativa de todos os trajectos possíveis pelo que, nestes casos, a modelização por trajectos candidatos consegue ser mais eficiente num número significativo de casos. No entanto, isto acontece à custa de tempos de cálculo bastante superiores. A conclusão que se retira é que apenas para redes pequenas (aquelas para as quais os tempos de cálculo absolutos não são significativos), ambos os algoritmos são aplicáveis desde que se considere um número significativo de trajectos candidatos.

Como conclusão, o algoritmo baseado em modelização explícita de VPCs é melhor do que o algoritmo baseado em trajectos candidatos pelo que deste ponto em diante apenas os resultados computacionais deste algoritmo serão considerados.

3.5.3 Dimensionamento Uni-Horário vs. Multi-Horário

O próximo aspecto a analisar é a avaliação dos ganhos que se obtêm no dimensionamento das redes quando estas permitem a reconfiguração dinâmica de VPCs. A Tabela 3-6 apresenta os valores mínimo, médio e máximo dos ganhos obtidos no cenário multi-horário comparativamente com o cenário uni-horário para todas as redes consideradas. Os resultados mostram que as redes dimensionadas num cenário multi-horário têm um custo menor cuja média varia de rede para rede mas que atinge em muitos casos valores bastante significativos.

Rede	Mínimo	Médio	Máximo	Rede	Mínimo	Médio	Máximo
10S	3,69%	9,70%	15,70%	10D	2,10%	8,08%	15,65%
15S	7,04%	11,79%	16,43%	15D	8,37%	13,78%	20,07%
20S	10,75%	15,04%	18,37%	20D	10,39%	21,58%	26,12%
25S	19,34%	20,51%	22,68%	25D	11,97%	15,84%	20,64%
30S	20,70%	22,65%	24,71%	30D	11,55%	13,53%	14,55%
35S	23,38%	24,25%	25,19%	35D	16,34%	17,63%	18,60%
40S	24,14%	24,99%	26,76%	40D	16,46%	18,75%	21,42%
45S	25,42%	26,30%	27,54%	45D	17,82%	21,19%	22,64%
50S	25,59%	26,20%	27,35%	50D	20,89%	21,95%	22,77%

Tabela 3-6 : Ganhos no dimensionamento das redes para o cenário multi-horário

Os resultados apresentados devem no entanto ser interpretados como uma ilustração da utilidade em se considerarem modelos de dimensionamento multi-horário. Os ganhos que se obtêm são fortemente dependentes de quão diferentes são os perfis de tráfego dos diferentes serviços suportados. Nos casos reais, estes modelos são úteis se a rede suportar serviços com diferentes perfis de tráfego, por exemplo, serviços cujos clientes alvo são as empresas e serviços cujos clientes alvo são as residências. Neste caso, é previsível que os dois tipos de serviços tenham períodos de maior tráfego em alturas diferentes do dia (os serviços empresariais com períodos de maior tráfego de manhã e à tarde e os serviços residenciais com períodos de maior tráfego à noite) e a reconfiguração dinâmica das redes lógicas de suporte a cada serviço permite efectivamente conseguir que uma rede de menor custo suporte o mesmo tráfego esperado.

3.5.4 Análise do Desempenho do Algoritmo

O próximo aspecto a considerar é a avaliação do algoritmo proposto em termos da sua eficiência computacional (tempos de cálculo) e eficácia (qualidade das soluções obtidas). En relação à eficiência, a Figura 3-2 apresenta os valores médios dos tempos de cálculo dos 10 problemas para cada rede e em cada caso horário.

Os resultados foram obtidos numa plataforma PC com um processador Pentium II a 350 MHz e com 256 Mbytes de memória RAM. O algoritmo mostrou-se bastante eficiente pois todas as redes foram dimensionadas em menos de 7 minutos numa plataforma computacional de baixo custo (note-se que o pior caso considera um conjunto de três redes lógicas com 780 VPCs em cada período horário num total de 7020 VPCs para todos os serviços e todos os períodos horários). Estes resultados mostram também que a complexidade computacional do algoritmo proposto é linear com o número de períodos

horários (os tempos de cálculo dos problemas multi-horário são cerca de três vezes maiores que os tempos de cálculo dos problemas uni-horário). Finalmente, é possível concluir também que a densidade de arestas em cada rede tem um impacto não negligenciável no tempo de resolução do algoritmo: as redes densas têm um número de arestas 66% maior do que as redes esparsas e o seu tempo de cálculo foi em média 30.3% superior.

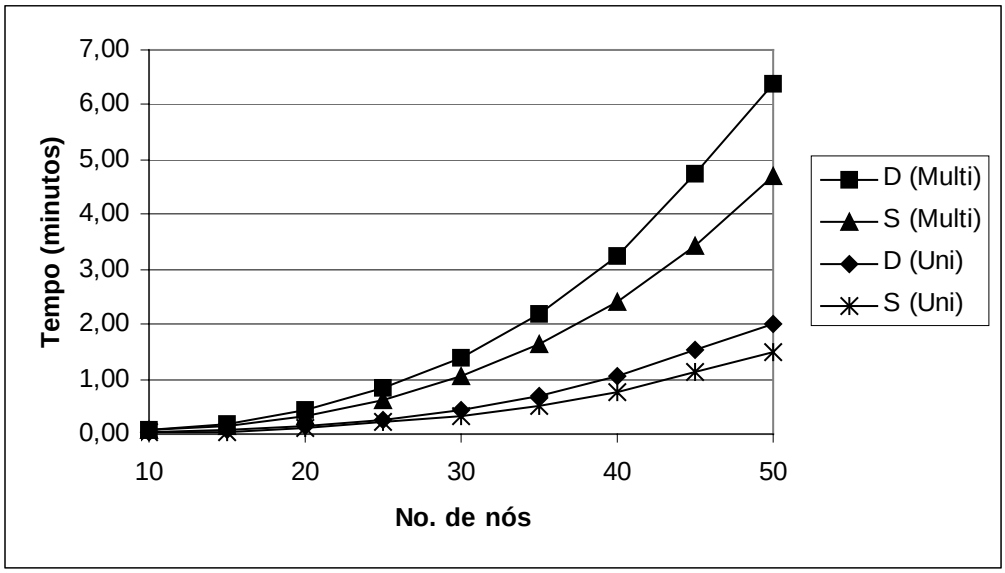


Figura 3-2 : Tempos de cálculo para as diferentes redes em ambos os casos horários

Em relação à eficácia do algoritmo, o intervalo dual de cada problema (dado pela percentagem em excesso do valor da melhor solução relativo ao limite inferior encontrado) permite-nos fazer uma análise da qualidade das soluções encontradas. A Figura 3-3 apresenta os valores mínimo e máximo dos intervalos duais obtidos pelo algoritmo nos 10 problemas uni-horário de cada rede. A Figura 3-4 apresenta a mesma informação para o caso multi-horário. Os valores destas figuras são apresentados na Tabela 3-7.

Note-se que um intervalo dual de X% significa que o custo da solução encontrada pelo algoritmo é no máximo X% superior ao custo da solução óptima do problema. Intervalos duais pequenos significam que a solução encontrada é de boa qualidade. Intervalos duais grandes significam que não é possível aferir da qualidade da solução encontrada pois não se sabe se o valor da solução óptima está mais próximo do limite inferior ou do custo da melhor solução encontrada. Assim, o intervalo dual serve como uma medida qualitativa da confiança que se tem na solução encontrada.

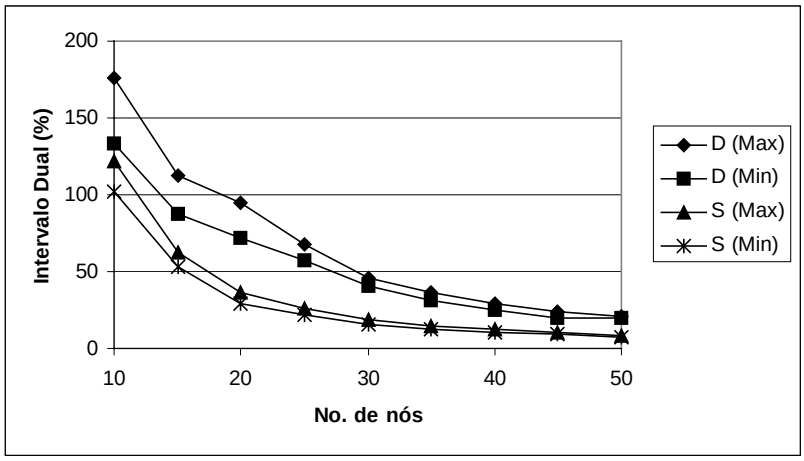


Figura 3-3 : Intervalo dual mínimo e máximo dos 10 problemas uni-horário para cada rede

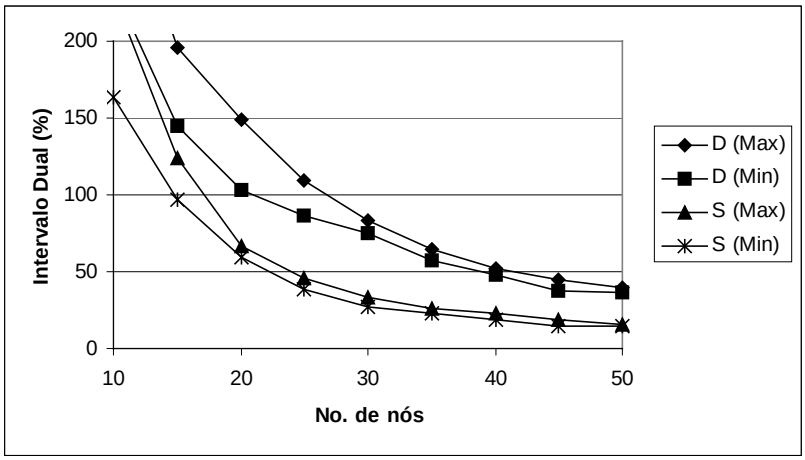


Figura 3-4 : Intervalo dual mínimo e máximo dos 10 problemas multi-horário para cada rede

Uni-horário				Multi-horário			
S (Min)	S (Max)	D (Min)	D (Max)	S (Min)	S (Max)	D (Min)	D (Max)
102,2	121,4	132,9	176,1	163,6	229,7	235,3	318,5
52,7	62,2	87,4	112,7	97,2	124,2	144,5	196,2
28,8	36,2	72,2	95,1	59,4	67,0	102,7	149,0
21,4	26,1	57,7	67,4	38,9	45,5	86,6	109,5
15,2	18,5	40,8	46,3	27,4	33,6	75,3	83,3
12,2	14,8	31,7	36,6	23,2	26,0	57,6	64,3
10,7	12,0	25,1	28,8	18,4	22,7	47,7	51,6
9,0	10,3	20,3	23,6	14,5	18,4	37,9	45,3
7,5	8,5	19,4	21,1	14,8	16,0	36,0	39,2

Tabela 3-7 : Intervalo dual mínimo e máximo dos 10 problemas para cada rede e cada caso horário

Os resultados mostram que a confiança nos resultados obtidos é maior para as redes esparsas do que para as redes densas e é também maior para os problemas uni-horário do que para os problemas multi-horário. Os resultados são particularmente bons para as redes

grandes e esparsas e não temos conhecimento que na literatura tenha sido descrito nenhum método de dimensionamento de redes que para estas dimensões consiga calcular em tempos semelhantes soluções com intervalos duais da ordem dos 10% como os obtidos pelo algoritmo proposto. Os resultados apresentados sugerem também que a confiança das soluções encontradas aumenta com a dimensão da rede. Esta conclusão é apenas parcialmente correcta como será visto mais adiante.

O valor da solução da relaxação Lagrangeana aplicada às restrições de capacidade do modelo original, conforme foi proposto na secção 3.4.2, é no máximo igual ao valor da relaxação linear do problema original, quaisquer que sejam os valores dos multiplicadores de Lagrange considerados. Esta conclusão é uma consequência do facto da relaxação Lagrangeana satisfazer a propriedade da integralidade, ou seja, a sua relaxação linear tem uma solução óptima inteira [Geoffrion74]. Usando o algoritmo *simplex* implementado na ferramenta CPLEX, foram calculados os valores da relaxação linear de todos os problemas uni-horário, que são também limites inferiores válidos para os problemas considerados. Verificou-se que o limite inferior encontrado pelo algoritmo proposto é no máximo 0.5% inferior ao valor dado pelo CPLEX. Conclui-se portanto que o algoritmo proposto calcula um limite inferior bastante próximo do seu máximo teórico. A Tabela 3-8 apresenta os tempos médios de cálculo do limite inferior dos dois métodos nos 10 problemas de todas as redes (executámos o CPLEX no mesmo PC em que foi executado o algoritmo proposto).

Rede	CPLEX	A.P.	Rede	CPLEX	A.P.
10S	0,10	1,12	10D	0,18	1,32
15S	0,76	3,07	15D	1,25	3,78
20S	1,98	6,78	20D	3,70	8,58
25S	5,68	12,31	25D	8,23	15,92
30S	9,45	20,22	30D	16,91	26,29
35S	17,21	31,24	35D	32,21	41,17
40S	33,63	46,35	40D	60,20	63,11
45S	58,71	66,97	45D	100,52	91,55
50S	91,63	89,84	50D	161,06	120,78

Tabela 3-8 : Tempos de execução (em segundos) do cálculo do limite inferior para os problemas uni-horário via relaxação linear (CPLEX) e via o algoritmo proposto (A.P.)

Pode-se verificar que embora os tempos de cálculo dos dois métodos sejam de mesma ordem de grandeza, há uma tendência para o algoritmo proposto ser mais rápido do que o cálculo da relaxação linear para as redes com um maior número de nós. Este facto indicia que para redes maiores que 50 nós o algoritmo proposto deverá ser um método

mais eficiente para o cálculo deste limite inferior do que a resolução da relaxação linear através do algoritmo *simplex*.

No caso dos problemas multi-horário, não foi possível calcular a relaxação linear para as redes maiores usando o CPLEX. A Tabela 3-9 apresenta os tempos de execução do CPLEX e do algoritmo proposto para estes problemas. Os tempos de execução do CPLEX apresentados nesta tabela para as redes com mais de 20 nós são referentes ao cálculo de apenas um dos dez problemas considerados. Claramente, o tempo de cálculo do algoritmo *simplex* explode para o modelo multi-horário. Por exemplo, o CPLEX demorou quase 34 horas para calcular a relaxação linear do problema associado à rede 40D e demorou quase 10 horas para calcular a relaxação linear do problema associado à rede 45S. No entanto, o limite inferior encontrado pelo algoritmo proposto foi em média inferior 7.2% do valor dado pelo CPLEX para todos os problemas em que foi calculada a relaxação linear. Este resultado mostra que uma das razões pelas quais o grau de confiança das soluções para os problemas multi-horário é menor do que nos casos uni-horário é a incapacidade do limite inferior encontrado atingir valores muito próximos dos seus valores máximos teóricos.

Rede	CPLEX	A.P.	Rede	CPLEX	A.P.
10S	2,44	3,31	10D	9,61	3,84
15S	33,71	9,32	15D	168,20	11,76
20S	215,34	20,10	20D	1568,70	26,59
25S	1184,75	37,75	25D	6404,91	49,63
30S	2787,89	63,64	30D	19153,52	83,10
35S	6039,50	98,48	35D	51086,4	131,69
40S	13542,16	144,34	40D	121993,1	194,68
45S	35372,2	206,18	45D	-	284,38
50S	-	282,54	50D	-	383,88

Tabela 3-9 : Tempos de execução (em segundos) do cálculo do limite inferior para os problemas multi-horário via relaxação linear (CPLEX) e via o algoritmo proposto (A.P.)

O valor da relaxação linear do problema original é fortemente dependente da relação entre a capacidade dos VPCs e as capacidades das ligações físicas disponíveis. Para ilustrar esta afirmação, considere-se o exemplo simples de uma rede com apenas 2 nós e uma aresta entre eles. Considere-se que é possível usar ligações de 10 Mbps com um custo de 1 e ligações de 100 Mbps com um custo de 6. Considere-se que a rede deve suportar um conjunto de VPCs entre os 2 nós da rede com capacidade total de 130 Mbps. Facilmente se determina que a solução óptima do problema é instalar uma ligação de 100 Mbps e 3 ligações de 10 Mbps com um custo total de 9. Por outro lado, a relaxação linear do

problema permite que o número de ligações de cada tipo possa ser um valor real. Assim, a solução óptima da relaxação linear é instalar 1.3 ligações de 100 Mbps com um custo total de 7,8. Neste caso, a relaxação linear do problema dá um limite inferior que está 13.3% abaixo da solução óptima. Se agora considerarmos que o conjunto de VPCs a suportar tem uma capacidade total de 430 Mbps, facilmente se verifica que a solução óptima tem custo 27 enquanto que a solução da relaxação linear tem custo 25.8. Neste caso, o limite inferior está apenas a 4.4% do valor óptimo. Conclui-se portanto que numa situação de maior tráfego a relaxação linear deu um limite inferior melhor do que numa situação de menor tráfego.

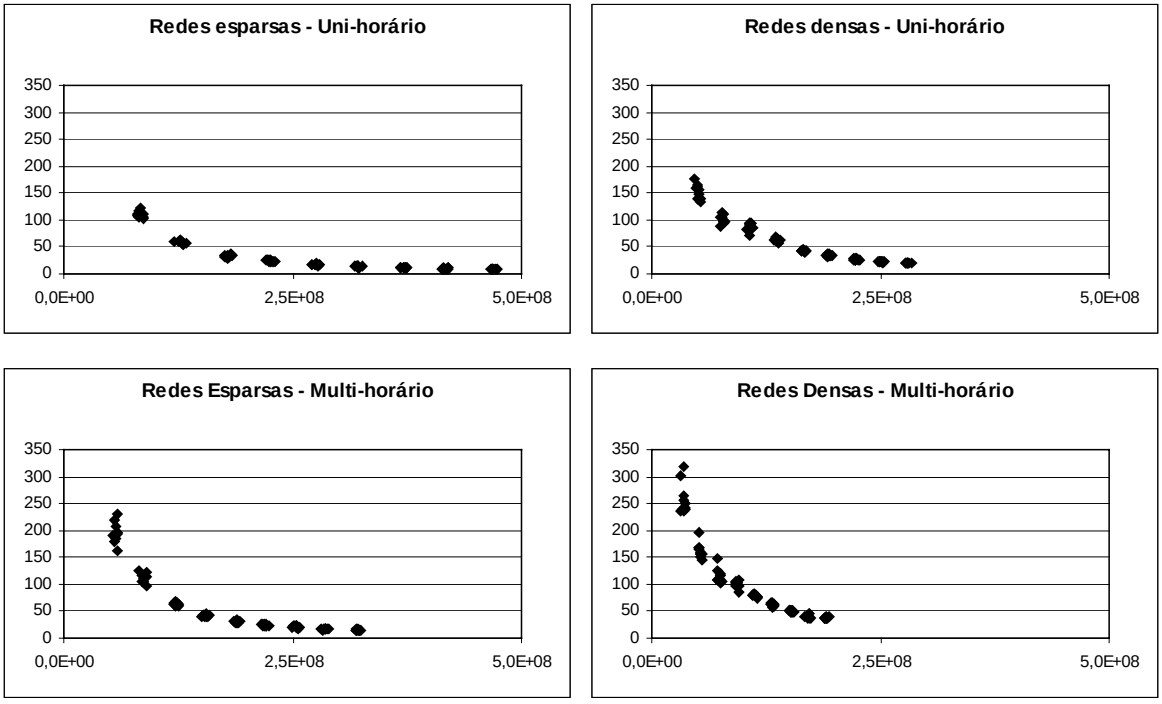


Figura 3-5 : Intervalo dual em função da capacidade média por aresta para os dois tipos de redes e nos dois casos horários

Este efeito também se verifica nos resultados computacionais obtidos para os problemas considerados. Para cada problema uni-horário, calculamos a capacidade média por aresta: a soma das capacidades de todos os VPCs a dividir pelo número de arestas da rede. No caso multi-horário, a capacidade média por aresta é dada pela soma das capacidades de todos os VPCs a dividir pelo número de períodos horários e pelo número de arestas. Com esta informação construimos os gráficos apresentados na Figura 3-5 em que cada problema é representado por um ponto cuja abcissa é a sua capacidade média por aresta e cuja ordenada é o seu intervalo dual. Os gráficos apresentados mostram que o

intervalo dual obtido é inversamente proporcional à capacidade média por aresta de cada problema. Verifica-se também que problemas com valores semelhantes de capacidade média por aresta exibem intervalos duais próximos em qualquer dos gráficos. Esta observação é mais evidente na Figura 3-6 em que os resultados anteriores são aglutinados num único gráfico.

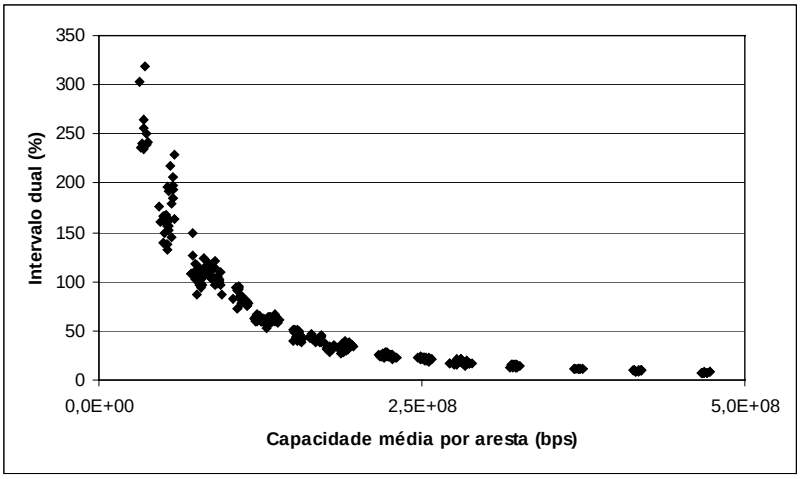


Figura 3-6 : Intervalo dual de cada problema em função da sua capacidade média por aresta

Estes resultados sugerem que a capacidade média por aresta de um problema de dimensionamento é o principal factor que determina o intervalo dual do algoritmo proposto, mais do que o número de nós, de arestas ou de VPCs. Segundo esta análise, qualquer que seja a topologia da rede, o intervalo dual pode ser arbitrariamente pequeno (grande) desde que para isso consideremos um tráfego total suficientemente grande (pequeno). Para confirmar esta análise, fizemos a experiência que a seguir se descreve. Usamos o algoritmo proposto para resolver o problema 1 das redes 10S e 10D (problemas com intervalos duais elevados) em que consideramos os valores das capacidades dos VPCs 10 vezes maiores que os originais. Usamos também o algoritmo proposto para a resolução do problema 1 das redes 50S e 50D (problemas com intervalos duais pequenos) em que consideramos os valores das capacidades dos VPCs 10 vezes menores que os originais. A Tabela 3-10 apresenta os resultados obtidos. Obtivemos desta forma uma grande redução dos intervalos duais dos problemas de 10 nós e um grande aumento dos intervalos duais dos problemas de 50 nós. A Figura 3-7 apresenta o gráfico apresentado anteriormente com a informação acrescida dos últimos problemas considerados. Estes resultados reforçam as conclusões apresentadas pois estes novos problemas têm uma relação entre a sua

capacidade média por aresta e o intervalo dual que apresentam, próxima dos problemas anteriormente considerados.

Uni-horário	10S	10D	50S	50D
Custo da melhor solução	15684,2	13948,6	12549,6	14058,8
Limite inferior	14547,7	11691,7	6704,2	4952,05
Intervalo dual (%)	7,81	19,30	87,19	183,90

Multi-horário	10S	10D	50S	50D
Custo da melhor solução	12879,8	11577,6	10400,3	9978,37
Limite inferior	10283,9	7913,6	4639,23	3389,64
Intervalo dual (%)	25,24	46,30	124,18	194,38

Tabela 3-10 : Resultados computacionais para os problemas alterados

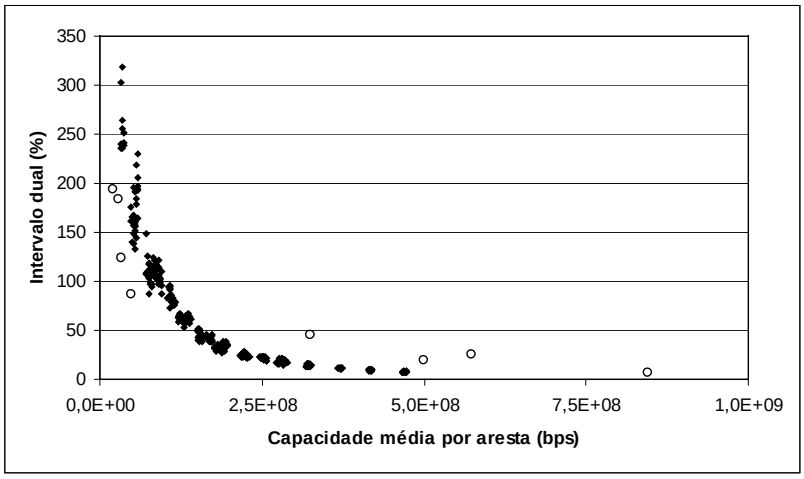


Figura 3-7 : Intervalo dual de cada problema em função da sua capacidade média por aresta

Estamos agora em condições para justificar a razão pela qual se observa nos resultados da Figura 3-3 e da Figura 3-4 que os intervalos duais decrescem (e por conseguinte a confiança das soluções encontradas aumenta) para as redes de maior dimensão. Note-se que, pela forma como as redes foram determinadas (secção 3.5.1), o número de arestas aumenta proporcionalmente com N (o número de nós da rede). Por outro lado, pela forma como os cenários de tráfego foram gerados (secção 3.5.1), a capacidade média de cada VPC é igual para todas as redes e o número de VPCs aumenta proporcionalmente com N^2-N. Assim, as redes de maior dimensão têm uma capacidade média por aresta maior e, por esta razão, exibem intervalos duais menores. Este é um cenário que consideramos previsível no crescimento das redes reais. Quando um operador acrescenta um novo nó na rede para ligar novos clientes, é previsível que o tráfego pelo

menos se mantenha entre todos os outros pares de nós de tráfego e que este novo nó gere tráfego adicional com todos os outros nós. Desta forma, o tráfego total aumenta percentualmente mais do que o número de nós da rede.

Em termos de desempenho, podemos assim concluir que o algoritmo proposto resolve eficientemente problemas de dimensionamento (tempos de cálculo pequenos) e a sua eficácia (grau de confiança na qualidade das soluções) depende fortemente da capacidade média por aresta do problema. Para problemas com intervalos duais grandes, não é possível saber a qualidade da solução encontrada a não ser que por outros métodos se consigam encontrar limites inferiores melhores dos que são cálculados pelo método proposto.

3.5.5 Análise do Limite Inferior do Problema

Uma forma alternativa de calcular limites inferiores para problemas de dimensionamento em grafos é através de restrições de corte da rede. Se dividirmos o conjunto de todos os nós da rede em dois sub-conjuntos disjuntos, então os VPCs cujos nós extremo não pertençam ao mesmo conjunto têm necessariamente de passar por uma das arestas do grafo cujos extremos também não pertençam ao mesmo conjunto. Então qualquer solução possível para o problema respeita uma restrição de capacidade que imponha que a soma das capacidades das ligações nas arestas entre nós que não pertençam ao mesmo sub-conjunto seja maior ou igual à soma das capacidades de todos os VPCs cujos nós extremo também não pertencem ao mesmo conjunto. Cada possível partição do conjunto de nós da rede em dois sub-conjuntos define uma restrição de corte. Então, um limite inferior válido para o problema de dimensionamento é dado pelo valor da solução do problema de optimização que considera a mesma função objectivo do problema original e que contém todas as restrições de corte possíveis. Este problema é formulável da seguinte forma:

Minimizar $$\sum_{\{i,j\}\in A}\sum_{t\in T} C^t_{\{ij\}} \cdot y^t_{\{ij\}}$$

Sujeito a:

$$\sum_{\{ij\}\in A_p}\sum_{t\in T} \alpha_t \cdot y^t_{\{ij\}} \geq M_p \qquad , p \in P$$

em que P é o conjunto de todas as partições do conjunto de nós da rede, A_p é o conjunto das arestas cujos extremos não pertencem ao mesmo conjunto na partição $p \in P$ e M_p é a soma das capacidades de todos os VPCs cujos extremos não pertencem ao mesmo conjunto na partição $p \in P$. Nos problemas com serviços assimétricos a soma das capacidades de todos os VPCs é diferente nos dois sentidos e o valor de M_p é dado pelo maior dos valores correspondentes a cada um dos sentidos. Nos problemas multi-horário, a soma das capacidades de todos os VPCs é diferente em cada período horário e o valor de M_p é dado pelo maior dos valores correspondentes a cada um dos períodos horários. O número total de restrições desta formulação (igual ao número total de partições de uma rede com N nós) é $2^{N-1}-1$. O facto da complexidade do modelo não aumentar para tráfego assimétrico nem para tráfego multi-horário é um dos principais méritos desta formulação.

Este problema foi resolvido através do algoritmo de *branch-and-bound* no CPLEX para todos os problemas associados às redes com 10 nós com tempos de execução de alguns minutos. Os limites inferiores obtidos pelo algoritmo de *branch-and-bound* aplicado ao modelo apresentado (referido como algoritmo alternativo) são apresentados na Tabela 3-11 para os problemas uni-horário e na Tabela 3-12 para os problemas multi-horário juntamente com os limites inferiores dados pelo algoritmo proposto baseado na relaxação Lagrangeana. Para as redes com 15 nós, o número de restrições do problema ($2^{14}-1=16383$) torna a sua resolução difícil com tempos de cálculo de várias horas. Por esta razão, este problema foi apenas resolvido para os dois primeiros problemas da rede 15S nos dois casos horários e os resultados (bem como os tempos de execução) são apresentados na Tabela 3-13.

	Rede 10S (Uni-horário)					**Rede 10D (Uni-horário)**				
	Custo	Algorit. Proposto		Algorit. Alternativo		Custo	Algorit. Proposto		Algorit. Alternativo	
		Limite	Int. Dual	Limite	Int. Dual		Limite	Int. Dual	Limite	Int. Dual
1	3012,6	1454,7	107,1%	2695,0	11,8%	2806,1	1169,2	140,0%	2508,6	11,9%
2	2967,7	1450,9	104,5%	2610,3	13,7%	2785,6	1063,1	162,0%	2467,5	12,9%
3	2900,0	1400,9	107,0%	2638,2	9,9%	2832,4	1136,6	149,2%	2539,1	11,6%
4	2966,4	1441,5	105,8%	2767,7	7,2%	2856,4	1034,5	176,1%	2445,6	16,8%
5	3030,3	1498,6	102,2%	2670,9	13,5%	2905,2	1127,0	157,8%	2528,4	14,9%
6	2942,2	1366,4	115,3%	2595,6	13,4%	2865,6	1201,4	138,5%	2521,6	13,6%
7	3186,0	1509,6	111,1%	2794,8	14,0%	2848,9	1196,6	138,1%	2506,6	13,7%
8	2851,1	1349,9	111,2%	2665,2	7,0%	2774,9	1064,3	160,7%	2376,1	16,8%
9	3042,0	1437,3	111,6%	2615,2	16,3%	2913,3	1251,0	132,9%	2574,5	13,2%
10	3164,0	1429,1	121,4%	2643,3	19,7%	2921,2	1098,2	166,0%	2493,3	17,2%
	Média:		109,7%		12,6%			152,1%		14,2%

Tabela 3-11 : Comparação dos intervalos duais pelos dois algoritmos nos problemas 10S uni-horário

	Rede 10S (Multi-horário)					Rede 10D (Multi-horário)				
	Custo	Algorit. Proposto		Algorit. Alternativo		Custo	Algorit. Proposto		Algorit. Alternativo	
		Limite	Int. Dual	Limite	Int. Dual		Limite	Int. Dual	Limite	Int. Dual
1	2754,4	968,8	184,3%	2378,3	15,8%	2367,0	703,9	236,3%	2049,8	15,5%
2	2835,2	859,9	229,7%	2345,5	20,9%	2484,0	740,9	235,3%	2000,8	24,2%
3	2698,1	925,1	191,7%	2278,1	18,4%	2755,4	756,6	264,2%	2015,4	36,7%
4	2856,8	973,2	193,5%	2426,3	17,7%	2599,2	645,1	302,9%	1981,4	31,2%
5	2875,7	938,7	206,3%	2333,0	23,3%	2831,0	676,5	318,5%	2189,4	29,3%
6	2480,2	940,9	163,6%	2278,1	8,9%	2653,3	782,2	239,2%	2093,9	26,7%
7	2759,3	928,6	197,1%	2376,6	16,1%	2789,0	816,8	241,4%	2182,5	27,8%
8	2561,9	918,1	179,0%	2278,1	12,5%	2426,8	714,2	239,8%	2061,1	17,7%
9	2611,8	820,7	218,3%	2307,5	13,2%	2778,7	792,0	250,8%	2231,7	24,5%
10	2694,0	916,9	193,8%	2440,4	10,4%	2529,4	710,2	256,2%	2014,5	25,6%
	Média:		195,7%		15,7%			258,5%		25,9%

Tabela 3-12 : Comparação dos intervalos duais pelos dois algoritmos nos problemas 10S multi-horário

	Rede 15S (Uni-horário)						Rede 15S (Multi-horário)					
	Custo	Alg. Proposto		Alg. Alternativo			Custo	Alg. Proposto		Alg. Alternativo		
		Limite	I. Dual	Limite	I. Dual	Tempo		Limite	I. Dual	Limite	I. Dual	Tempo
1	6248,1	3991,4	56,5%	5263,2	18,7%	15:46:30	5409,7	2744,0	97,2%	4427,3	22,2%	18:54:31
2	6367,4	3934,8	61,8%	5341,6	19,2%	40:52:09	5321,0	2578,3	106,4%	4532,2	17,4%	38:32:57

Tabela 3-13 : Comparação dos intervalos duais pelos dois algoritmos nos problemas 15S considerados

Estes resultados mostram que para os problemas considerados, o limite inferior dado pelo algoritmo baseado em Relaxação Lagrangeana é bastante mau. No caso das redes com 10 nós, os novos limites calculados pela formulação alternativa reduzem os intervalos duais anteriores para valores bastante inferiores mostrando assim que a solução encontrada pelo algoritmo proposto está bastante mais perto do valor óptimo de que o seu limite inferior. No caso dos dois problemas de dimensionamento das redes 15S, as conclusões são semelhantes embora a melhoria seja menor.

Os resultados indiciam claramente que a eficácia do algoritmo proposto é bastante maior do que os intervalos duais fazem parecer. Para os casos em que foi possível calcular um limite inferior pela formulação alternativa apresentada, foi possível confirmar que os limites inferiores calculados pelo algoritmo proposto são bastante maus. Para que estas conclusões possam ser extendidas às redes maiores, é no entanto necessário mais investigação sobre formas mais eficientes de calcular limites inferiores para o modelo estudado. [Magnanti95] e [Bienstock96] são duas referências que ilustram o estado-da-arte neste assunto. Nesses trabalhos, são propostas diferentes estratégias e diferentes classes de restrições para a reformulação de problemas de dimensionamento de redes que, embora

mais simples, têm muitas características comuns ao problema abordado nesta tese. Os resultados mostram que é possível fazer melhor do que a formulação alternativa aqui proposta mas a complexidade computacional dos vários métodos não permite endereçar problemas de dimensionamento muito para além de redes com 15 nós.

3.5.6 Observações Finais

Foram realizados cálculos computacionais tendo por base as redes propostas na secção 3.5.1 e considerando cenários de tráfego com os três tipos de serviços endereçados (conversacionais, de consulta e ponto-multiponto). As conclusões que se retiram dos resultados obtidos são bastante semelhantes às que se descrevaram nas sub-secção anteriores, razão pela qual esses resultados não são incluídos nesta tese. Uma conclusão que se retira dos resultados obtidos é que o desempenho do algoritmo é independente do facto dos VPCs serem ou não serem assimétricos. Em relação aos serviços de consulta, consideramos como caso de interesse apenas cenários em a rede suporta simultaneamente este tipo de serviços e os os outros tipos de serviços (conversacionais e/ou ponto-multiponto). Neste caso, os serviços de consulta não têm grande influência no desempenho global do algoritmo porque gera um número de VPCs bastante menor do que os outros tipos de serviços.

3.6 Conclusões

Neste capítulo foi abordado o problema do dimensionamento de redes ATM configuradas com redes lógicas exclusivamente constituídas por VPCs extremo-a-extremo. Foram apresentados dois modelos que definem os problemas de dimensionamento considerando o suporte a serviços conversacionais, serviços de consulta e serviços ponto-multiponto. Um modelo é baseado na modelização explícita de VPCs e o outro é baseado em trajectos candidatos. Foi proposto um algoritmo heurístico baseado na Relaxação Lagrangeana com Optimização por Sub-gradiente para a resolução dos modelos propostos. O algoritmo proposto mostrou-se bastante eficiente. Pelos resultados computacionais foi possível demonstrar que a modelização explícita de VPCs é mais eficiente do que a modelização por trajectos candidatos porque consegue obter soluções em média melhores com tempos de cálculo menores. Os resultados computacionais mostram também que os modelos multi-horário permitem ganhos no custo da rede física quando as redes permitem

reconfiguração dinâmica de VPCs. Pela análise dos resultados computacionais, verificou-se que o principal factor que determina a qualidade das soluções encontradas pelo algoritmo proposto é a capacidade média dos VPCs por aresta da rede. Finalmente, para as redes pequenas foi possível calcular limites inferiores teóricos da solução óptima bastante melhores do que os determinados pelo algoritmo proposto indiciando assim que a qualidade das soluções dadas pelo algoritmo proposto é melhor do que os intervalos duais que ele determina fazem parecer.

CAPÍTULO 4

REDES LÓGICAS HIERARQUIZADAS A DOIS NÍVEIS

Com o crescimento das redes ATM, uma arquitectura exclusivamente baseada em VPCs extremo-a-extremo tornar-se-á virtualmente impraticável devido aos problemas operacionais que um grande número de VPCs impõe na gestão de recursos. Um dos problemas é o tamanho das tabelas de encaminhamento nos comutadores ATM (exigindo grande capacidade de memória) e a degradação do desempenho das funções de comutação que se baseiam em operações de leitura destas tabelas. Outro aspecto importante são os procedimentos de recuperação a falhas que necessariamente têm de estar presentes na rede. Em caso de falha de um elemento da rede, estes procedimentos reencaminham os VPCs afectados por trajectos alternativos e o tempo médio de recuperação a falhas é directamente proporcional ao número médio de VPCs existente na rede.

A gestão de recursos baseada em redes lógicas exclusivamente constituídas por VPCs extremo-a-extremo não explora ganhos de partilha de recursos ao nível da ligação lógica dado que cada VPC suporta apenas o tráfego entre os seus nós extremo. Uma rede lógica de suporte a um serviço com uma arquitectura em que o tráfego entre os nós da rede percorra mais do que um VPC, conduz a que diferentes fluxos de tráfego possam partilhar VPCs comuns e, consequentemente, haja ganhos de partilha de recursos. Quanto menor for o número de VPCs de uma rede lógica, maiores serão os ganhos de partilha de recursos

dado que o número médio de classes de tráfego que atravessam cada VPC é maior. No entanto, as capacidades das ligações físicas têm valores modulares e no caso geral não é possível ter uma rede física em que as capacidades das ligações físicas estejam completamente ocupadas pelos VPCs que por elas passam. Há sempre alguma capacidade extra nas ligações físicas que não é usada pelos VPCs. Numa rede lógica com um menor número de VPCs, a capacidade média de cada VPC é normalmente maior do que no caso de VPCs extremo-a-extremo. Assim, o efeito da modularidade das ligações físicas faz com que a capacidade média não usada seja maior do que com VPCs extremo-a-extremo. No caso geral, a eficiência do dimensionamento da rede física com redes lógicas segundo uma arquitectura genérica é um compromisso entre estes dois factores. Uma arquitectura é mais eficiente do que outra se os ganhos de partilha de recursos compensarem a capacidade não usada introduzida pela modularidade das ligações físicas ou vice-versa.

Neste capítulo, é proposta uma arquitectura hierarquizada a 2 níveis para a topologia das redes lógicas. Nesta arquitectura é possível, com pequenas aproximações, combinar o cálculo das capacidades dos VPCs com o dimensionamento da rede física através de modelos de Programação Inteira. São propostos dois modelos para o dimensionamento das redes com base nesta arquitectura: um modelo uni-horário e um modelo multi-horário. Os modelos de dimensionamento propostos neste capítulo endereçam o suporte de serviços conversacionais simétricos.

Este capítulo está organizado da seguinte forma. Na secção 4.1 é descrita a arquitectura proposta bem como as suas características principais. As aproximações consideradas que permitem combinar o cálculo das capacidades dos VPCs com o dimensionamento são apresentadas na secção 4.2. A secção 4.3 descreve o processo de cálculo das capacidades dos VPCs. Os modelos de optimização que enunciam o problema do dimensionamento da rede física são apresentados nas secções 4.4 (modelo uni-horário) e 4.5 (modelo multi-horário). Na secção 4.6 são apresentados e discutidos os resultados computacionais.

4.1 Arquitectura das Redes Lógicas

A arquitectura hierarquizada a 2 níveis é ilustrada na Figura 4-1. Os nós da rede são agrupados em regiões distintas. Em cada região será seleccionado um nó, designado por nó fronteira, responsável por encaminhar todo o tráfego entre a sua região e qualquer uma das

outras. Os outros nós de cada região são designados por nós interiores. Dentro de cada região, são configurados VPCs entre todos os pares de nós: este é o nível hierárquico inferior. No nível hierárquico superior, são configurados VPCs entre todos os pares de nós fronteira.

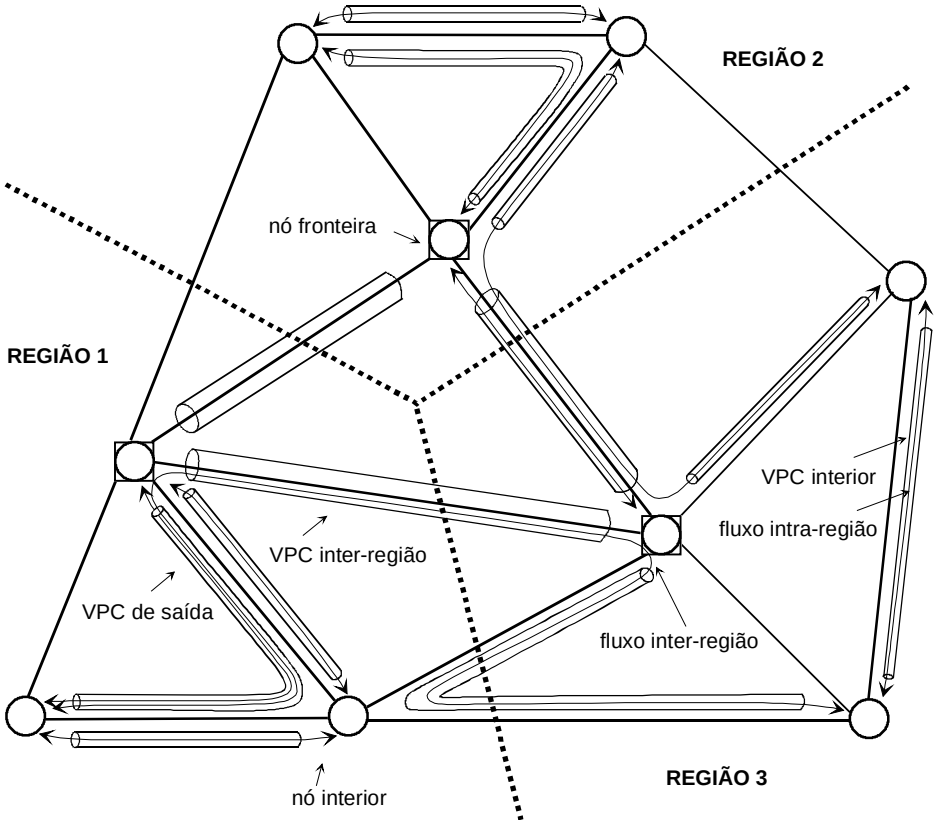


Figura 4-1 : Ilustração de uma rede lógica hierarquizada a 2 níveis

Os VPCs do nível hierárquico superior são designados por VPCs inter-região. Os outros são designados por VPCs intra-região. De entre estes últimos, distinguimos os que ligam nós interiores entre si, designados por VPCs interiores, e os que ligam um nó interior e o nó fronteira, designados por VPCs de saída. Analogamente, designamos as classes de tráfego entre nós da mesma região por classes de tráfego intra-região e as classes de tráfego entre nós de diferentes regiões por classes de tráfego inter-região. Os VPCs interiores suportam apenas o tráfego entre os seus nós extremo (são VPCs extremo-a-extremo). Os VPCs de saída suportam o tráfego entre os seus nós extremo e o tráfego entre o nó interior e todos os nós das outras regiões. Os VPCs inter-região suportam o tráfego entre os nós das regiões que interligam. Nesta arquitectura, existem ganhos significativos de partilha de recursos nos VPCs inter-região e nos VPCs de saída de cada região.

A arquitectura proposta apresenta alguns aspectos importantes. O primeiro aspecto é a redução do número de VPCs da rede. Por exemplo, uma rede com 50 nós de tráfego para um serviço precisa de uma rede lógica de suporte constituída por 1225 VPCs no caso de VPCs extremo-a-extremo enquanto que numa arquitectura hierarquizada com 5 regiões de 10 nós em cada região, necessita apenas de 235 VPCs o que equivale a uma redução de cerca de 80% no número total de VPCs. Outra característica importante é o facto de uma ligação lógica atravessar no máximo 3 VPCs. O tempo de estabelecimento de ligação é um parâmetro de qualidade de serviço ao nível da ligação lógica que não foi considerado no capítulo anterior porque com VPCs extremo-a-extremo este parâmetro é reduzido ao seu valor mínimo. Em arquitecturas em que as ligações lógicas atravessam múltiplos VPCs, este parâmetro deve ser tido em conta. Na arquitectura proposta, o facto de cada ligação ser estabelecida no máximo por 3 VPCs resulta numa degradação insignificante do tempo de estabelecimento de ligação. Finalmente, esta arquitectura conduz a redes físicas com características mais próximas do modo de funcionamento de um operador de telecomunicações. Os operadores usam habitualmente dois tipos de recursos dependendo do tipo e capacidade do equipamento de comutação que necessitam instalar em cada nó de rede. Para nós de comutação com pouco tráfego, os operadores usam cabines de rua sem nenhuma infra-estrutura de apoio residente. Nós de comutação com muito tráfego são habitualmente instalados em edifícios com pessoal permanente de operação e manutenção. Uma arquitectura hierarquizada a dois níveis resulta em redes físicas com nós fronteira de grande capacidade de comutação e nós interiores de pequena capacidade de comutação. No caso de redes lógicas constituídas exclusivamente por VPCs extremo-a-extremo, as redes resultantes têm tendência a ter todos os nós com níveis semelhantes de capacidade de comutação, o que dificulta a escolha mais apropriada do tipo de recursos necessários instalar pelo operador.

4.2 Aproximações Consideradas

Note-se que na arquitectura proposta, se os nós fronteira forem especificados à partida, então saberiamos quais os nós extremo de todos os VPCs e que classes de tráfego passam em cada um dos VPCs. Desta forma, a capacidade dos VPCs seria independente do seu trajecto e poderia ser calculada de uma forma independente. Se este fosse o caso, o modelo de dimensionamento resultante seria idêntico ao apresentado no capítulo anterior

para serviços conversacionais. A complexidade do problema aqui proposto advém do facto de querermos usar a escolha do nó fronteira de cada região em proveito do dimensionamento da rede física. A capacidade dos VPCs depende da localização dos nós fronteira em cada região pelo que não é possível separar o seu cálculo do dimensionamento da rede física. Para se conseguir combinar o cálculo das capacidades dos VPCs com o dimensionamento da rede (que determina o trajecto dos VPCs) através de um modelo de Programação Inteira, consideraram-se duas aproximações.

Na primeira aproximação, assume-se que em cada VPC de saída não existe partilha de recursos entre a classe de tráfego intra-região e o conjunto das classes de tráfego inter-região que o atravessam. Com esta aproximação, a capacidade de um VPC de saída é determinada pela soma da capacidade necessária para suportar a classe de tráfego intra-região com a capacidade necessária para o conjunto das classes de tráfego inter-região. Esta aproximação é conservativa porque a "capacidade real" necessária é menor ou igual à soma destas duas capacidades. Esta aproximação é boa porque a "capacidade real" é maior ou igual do que cada uma destas capacidades e a capacidade para suportar as classes inter-região é, em geral, muito maior que a capacidade para suportar a classe intra-região.

A segunda aproximação consiste em considerar que todas as classes de tráfego inter-região atravessam 3 VPCs. Esta aproximação equivale a considerar que todos os nós de uma região são interiores (mesmo o nó fronteira a escolher posteriormente). Com esta aproximação, o cálculo das capacidades dos VPCs inter-região e das capacidades dos VPCs de saída que suportam o tráfego inter-região torna-se independente da escolha do nó fronteira de cada região. Esta aproximação é conservativa porque estamos a considerar que algumas das classes de tráfego inter-região atravessam mais VPCs do que na realidade acontece pelo que a sua probabilidade de bloqueio real é menor do que a determinada pela aproximação. Esta aproximação é tanto mais melhor quanto maior for o número de nós de cada região.

Estas duas aproximações permitem, para efeitos de cálculo de capacidades dos VPCs, separar cada rede lógica em duas redes: uma rede de suporte às classes de tráfego intra-região (Figura 4-2a) e uma rede de suporte às classes de tráfego inter-região (Figura 4-2b). A Figura 4-2 ilustra esta separação no exemplo apresentado na Figura 4-1. A rede de suporte ao tráfego intra-região é constituída por VPCs extremo-a-extremo e, portanto, as

suas capacidades são calculadas tal como descrito na secção 3.1.1. As capacidades da rede de suporte ao tráfego inter-região são calculadas através da aproximação de carga reduzida.

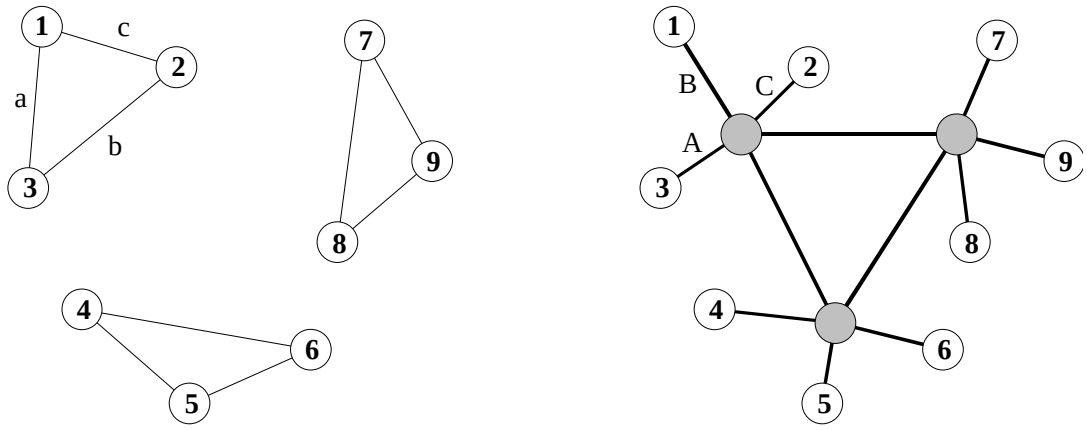


Figura 4-2 : Separação de uma rede lógica para classes de tráfego intra-região (a) e inter-região (b)

Depois de calculados todos os valores para estas duas redes, é possível calcular todos os valores que os VPCs inter-região e intra-região podem assumir em função da escolha do nó fronteira. Segundo as aproximações consideradas, a capacidade dos VPCs inter-região é independente do nó fronteira e é directamente dada pelo cálculo das capacidades de suporte ao tráfego inter-região (Figura 4-2b).

No caso dos VPCs intra-região, existem 3 casos possíveis. Para os ilustrar, considere os dois conjuntos de capacidades calculadas para a região composta pelos nós 1, 2 e 3 representadas pelas letras apresentadas na Figura 4-2. Por exemplo, *b* é a capacidade que é necessário atribuir ao VPC intra-região entre os nós 2 e 3 para suportar o tráfego entre estes dois nós; *A* é a capacidade necessária para suportar o tráfego entre o nó 3 e os nós de todas as outras regiões e *C* é a capacidade necessária para suportar o tráfego entre o nó 2 e os nós de todas as outras regiões. Então, a capacidade do VPC intra-região entre os nós 2 e 3 é:

- *b* se nem o nó 2 nem o nó 3 forem o nó fronteira da região
- *b* + *C* se o nó 3 for o nó fronteira da região (Figura 4-3a)
- *b* + *A* se o nó 2 for o nó fronteira da região (Figura 4-3b)

Estes são os 3 valores de capacidade que um VPC intra-região pode assumir. O primeiro caso corresponde à situação em que o VPC é interior. Os dois outros casos correspondem às duas situações em que o VPC é de saída.

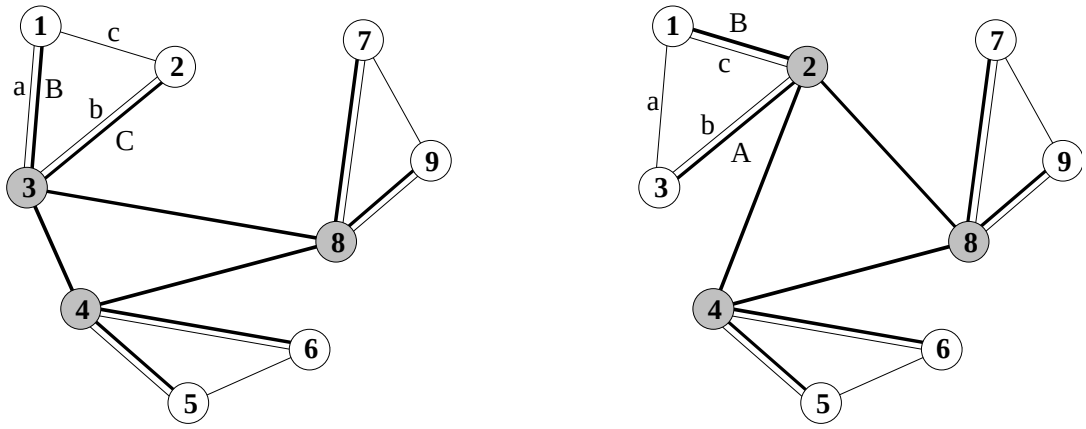


Figura 4-3 : Ilustração do cálculo das capacidades dos VPCs intra-região

Assim, a adopção da topologia hierarquizada a dois níveis e as aproximações introduzidas permitem reduzir a complexidade entre o cálculo das capacidades e o dimensionamento da rede pois existe um valor único de capacidade para os VPCs inter-região e um conjunto de três valores possíveis para os VPCs intra-região, que podem ser considerados no problema de dimensionamento da rede através de um conjunto de restrições lineares no respectivo modelo de optimização.

4.3 Cálculo das Capacidades dos VPCs

Considere-se o conjunto de nós da rede N, o conjunto de regiões R e o conjunto de nós pertencentes à região r, N_r. Considere-se o conjunto de serviços S suportados pela rede. A capacidade das ligações lógicas do serviço s é dada por w_s. O conjunto das classes de tráfego do serviço s é dado por F_s e cada classe de tráfego f é definido pelo seu tráfego ρ_f^s e pelo seu conjunto de nós extremo E_f^s.

A rede lógica de suporte ao serviço s é constituída pelo conjunto de VPCs inter-região L_s e pelo conjunto de VPCs intra-região K_s. Cada VPC inter-região $l \in L_s$ é definido por uma região origem o_l^s, uma região destino d_l^s e uma capacidade b_l^s. Cada VPC intra-região $k \in K_s$ é definido por um nó origem o_k^s, um nó destino d_k^s e um de três possíveis valores de capacidade b_k^{sp} (p = 0, 1, 2). O valor com índice p = 0 representa a capacidade quando o VPC k é interior; o valor com índice p = 1 representa a capacidade quando o

VPC *k* tem origem no nó fronteira da sua região e valor com índice *p* = 2 representa a capacidade quando o VPC *k* tem destino no nó fronteira da sua região.

Para calcular os valores de b_l^s e de b_k^{sp}, é preciso primeiro calcular os valores para a rede de suporte ao tráfego inter-região (Figura 4-2b). Assim, considere-se que $\tilde{i}$ designa o VPC de saída do nó *i* para o nó fronteira. Seja $\overline{f}$ o VPC inter-região atravessado pela classe de tráfego *f*. O conjunto de VPCs que definem o trajecto de uma classe de tráfego inter-região é dado por:

$$R_f^s = \{ \tilde{i} , \overline{f} , \tilde{j} \} \qquad , i \in E_f^s , j \in E_f^s - \{i\} \qquad (1)$$

Considere-se F_v^s como sendo o conjunto de classes de tráfego que atravessam o VPC *v*. Considere-se a variável m_v^s que especifica o número máximo de ligações suportado pelo VPC *v*. Pela aproximação de carga reduzida [Whitt85], a probabilidade de bloqueio de cada classe de tráfego inter-região *f*, B_f^s é dada por:

$$B_f^s = 1 - \left(1 - L_{\overline{f}}^s\right)\left(1 - L_{\tilde{i}}^s\right)\left(1 - L_{\tilde{j}}^s\right) \qquad , s \in S,\ f \in F_s, i \in E_f^s ,\ j \in E_f^s - \{i\} \quad (2a)$$

em que os valores L_v^s são dados pelo seguinte conjunto de equações de ponto fixo (ER[ρ,*m*] representa a formula de ErlangB):

$$L_l^s = ER\left[\sum_{f \in F_l^s} \rho_f^s \left(1 - L_{\tilde{i}}^s\right)\left(1 - L_{\tilde{j}}^s\right),\ m_l^s \right] \qquad , s \in S, l \in L_s, i \in E_f^s ,\ j \in E_f^s - \{i\} \quad (2b)$$

$$L_{\tilde{i}}^s = ER\left[\sum_{f \in F_{\tilde{i}}^s} \rho_f^s \left(1 - L_{\overline{f}}^s\right)\left(1 - L_{\tilde{j}}^s\right), m_{\tilde{i}}^s \right] \qquad , s \in S, i \in N,\ j \in E_f^s - \{i\} \qquad (2c)$$

Este conjunto de equações (2) pode ser resolvido através do método de substituições repetitivas [Ross95] (proposto em [Whitt85]) e permite calcular as probabilidades de bloqueio de todas as classes de tráfego dados os seus valores de tráfego e as capacidades m_v^s (em número máximo de ligações) de cada VPC. O nosso objectivo é fazer o cálculo inverso, ou seja, calcular o valor mínimo das capacidades dos VPCs de modo que todas as classes de tráfego tenham uma probabilidade de bloqueio menor ou igual que o valor máximo permitido para cada serviço. Este cálculo não tem uma solução única. O algoritmo adoptado é o que a seguir se descreve.

Inicialmente, os valores de m_v^s são calculados com base no teorema do limite do produto [Whitt85]. Considerando um valor de L_s dado por:

$$L_s = 1 - \sqrt[3]{1 - \Gamma_s} \tag{3}$$

em que Γ_s representa a probabilidade de bloqueio máxima permitida para o serviço *s*, calculam-se os menores valores de m_v^s que cumprem com as seguintes desigualdades:

$$ER\left[\sum_{f \in F_l^s} \rho_f^s \ , m_l^s\right] \leq L_s \qquad , s \in S, l \in L_s \tag{4a}$$

$$ER\left[\sum_{f \in F_{\tilde{i}}^s} \rho_f^s , m_{\tilde{i}}^s\right] \leq L_s \qquad , s \in S, i \in N \tag{4b}$$

As desigualdades (4) garantem que para os valores de m_v^s escolhidos, os valores de L_v^s, dados por (2b) e (2c) são menores que L_s e a expressão (3) garante que as probabilidades de bloqueio B_f^s de todas as classes de tráfego, dadas por (2a) são menores que Γ_s.

De seguida, tenta-se diminuir os valores m_v^s verificando se as probabilidades de bloqueio de todas as classes de tráfego permanecem menores que os seus valores máximos. Este processo é implementado da seguinte forma: escolhe-se um VPC *v*, reduz-se o valor de m_v^s em uma unidade e verifica-se se com este novo valor todas as classes de tráfego têm uma probabilidade de bloqueio menor que Γ_s (resolvendo as equações (2) pelo método de substituições repetitivas). Se sim, considera-se o novo valor de m_v^s e se não, retoma-se o valor anterior. Percorrem-se sequencialmente todos os VPC *v* e verifica-se para cada um se é possível reduzir o valor de m_v^s. O algoritmo termina no primeiro ciclo em que todos os VPCs sejam testados sem que tenha sido possível reduzir o valor de m_v^s em nenhum deles.

Depois de calculadas as variáveis m_l^s e $m_{\tilde{i}}^s$, determinam-se os valores das capacidades dos VPCs conforme descrito na secção 4.2. As capacidade dos VPCs inter-região $l \in L_s$ são dadas por:

$$b_l^s = w_s \cdot m_l^s \qquad , s \in S \, , l \in L_s \tag{5a}$$

e as capacidades dos VPCs intra-região são dadas por:

$$b_k^{s1} = b_k^{s0} + w_s \cdot m_{\tilde{i}}^s \qquad , i = d_k^s \tag{5b}$$

$$b_k^{s2} = b_k^{s0} + w_s \cdot m_{\tilde{i}}^s \qquad , i = o_k^s \tag{5c}$$

em que o valor de b_k^{s0} é calculado com base na fórmula de ErlangB tal como descrito na secção 3.1.1.

4.4 Modelo de Dimensionamento Uni-Horário

Depois de calculados todos os valores de capacidade dos VPCs inter-região (b_l^s) e VPCs intra-região (b_k^{sp}), o modelo de programação inteira seguinte formula o problema de optimização cuja resolução escolhe a combinação de trajectos para cada VPC (dados pelas variáveis x_{ij}^{skp} e x_{ij}^{sl}) e os nós fronteira (dados pelas variáveis z_i^{sr}) que resultam na rede física de menor custo:

N Conjunto de nós da rede

R Conjunto de regiões r da rede

N_r Conjunto dos nós de rede pertencentes à região $r \in R \left(N = \bigcup_r N_r \right)$

A Conjunto de arestas $\{i,j\}$ da rede: (i,j) representa a aresta $\{i,j\}$ no sentido de i para j e (j,i) representa o sentido inverso

T Conjunto dos tipos t de ligações físicas possíveis de implementar entre os nós de rede

α_t Largura de banda de uma ligação física do tipo t

$Y_{\{ij\}}^t$ Número máximo de ligações físicas do tipo t entre os nós i e j

$C_{\{ij\}}^t$ Custo associado à utilização de uma ligação física do tipo t entre os nós i e j

S Conjunto de serviços s suportados pela rede

K_s Conjunto de VPCs intra-região k da rede lógica de suporte ao serviço s; cada VPC k é caracterizado por:

- o_k^s Nó origem
- d_k^s Nó destino
- r_k^s Região à qual pertence ($r_k^s \in R$)
- b_k^{sp} Para $p = 0$, capacidade quando o VPC é interior; para $p = 1$, capacidade quando o_k^s é o nó fronteira de r_k^s; para $p = 2$, capacidade quando d_k^s é o nó fronteira de r_k^s

L_s Conjunto de VPCs inter-região l da rede lógica de suporte ao serviço s; cada VPC l é caracterizado por:

o_l^s Região origem ($o_l^s \in R$)

d_l^s Região destino ($d_l^s \in R$)

b_l^s Capacidade

$y_{\{ij\}}^t$ Variável inteira que determina o número de ligações físicas do tipo *t* entre os nós *i* e *j*

x_{ij}^{sl} Variável binária de trajecto: o valor 1 significa que o VPC inter-região *l* de suporte ao serviço *s* passa no arco (*i*,*j*)

x_{ij}^{skp} Variável binária de trajecto: para $p = 0$, o valor 1 significa que o VPC intra-região *k* é interior e passa no arco (*i*,*j*); para $p = 1$, o valor 1 significa que o_k^s é o nó fronteira de r_k^s e o VPC passa no arco (*i*,*j*); para $p = 2$, o valor 1 significa que d_k^s é o nó fronteira de r_k^s e o VPC passa no arco (*i*,*j*)

z_i^{sr} Variável binária: o valor 1 significa que o nó $i \in N_r$ é o nó fronteira da região $r \in R$ na rede lógica de suporte ao serviço *s*

Minimizar $$\sum_{\{i,j\}\in A}\sum_{t\in T} C_{\{ij\}}^t y_{\{ij\}}^t \qquad (6a)$$

Sujeito a:

$$\sum_{j\in N:\{i,j\}\in A}\sum_{p=0}^{2}\left(x_{ij}^{skp} - x_{ji}^{skp}\right)=1 \qquad , s \in S\,,\ k \in K_s\,,\ i = o_k^s \qquad (6b)$$

$$\sum_{j\in N:\{i,j\}\in A}\sum_{p=0}^{2}\left(x_{ji}^{skp} - x_{ij}^{skp}\right)=1 \qquad , s \in S\,,\ k \in K_s\,,\ i = d_k^s \qquad (6c)$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{ij}^{skp} - x_{ji}^{skp}\right)=0 \qquad , s \in S\,,\ k \in K_s\,,\ \text{todo } p\,,\ i \in N \setminus \{ o_k^s\,, d_k^s \} \qquad (6d)$$

$$\sum_{i\in N_r}\ \sum_{j\in N:\{i,j\}\in A}\left(x_{ij}^{sl} - x_{ji}^{sl}\right)=1 \qquad , s \in S\,,\ l \in L_s\,,\ r = o_l^s \qquad (6e)$$

$$\sum_{i\in N_r}\ \sum_{j\in N:\{i,j\}\in A}\left(x_{ji}^{sl} - x_{ij}^{sl}\right)=1 \qquad , s \in S\,,\ l \in L_s\,,\ r = d_l^s \qquad (6f)$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{ij}^{sl} - x_{ji}^{sl}\right)=0 \qquad , s \in S\,,\ l \in L_s\,,\ i \in N \setminus \left\{N_{o_l^s}\,, N_{d_l^s}\right\} \qquad (6g)$$

$$z_i^{sr} = \sum_{j\in N:\{i,j\}\in A}\left(x_{ij}^{sl} - x_{ji}^{sl}\right) \qquad , s \in S\,,\ r \in R\ ,\ i \in N_r\,,\ l \in L_s\colon\ o_l^s = r \qquad (6h)$$

$$z_i^{sr} = \sum_{j\in N:\{i,j\}\in A}\left(x_{ji}^{sl} - x_{ij}^{sl}\right) \qquad , s \in S\,,\ r \in R\ ,\ i \in N_r\,,\ l \in L_s\colon\ d_l^s = r \qquad (6i)$$

$$\sum_{j \in N:\{i,j\} \in A} \left(x_{ij}^{skp} - x_{ji}^{skp} \right) = z_i^{sr_k} \qquad , s \in S \, , k \in K_s \, , i = o_k^s \, , p = 1 \tag{6j}$$

$$\sum_{j \in N:\{i,j\} \in A} \left(x_{ji}^{skp} - x_{ij}^{skp} \right) = z_i^{sr_k} \qquad , s \in S \, , k \in K_s \, , i = d_k^s \, , p = 2 \tag{6k}$$

$$\sum_{s \in S} \sum_{k \in K_s} \sum_{p=0}^{2} b_k^{sp} \left(x_{ij}^{skp} + x_{ji}^{skp} \right) + \sum_{s \in S} \sum_{l \in L_s} b_l^s \left(x_{ij}^{sl} + x_{ji}^{sl} \right) \leq \sum_{t \in T} \alpha_t y_{\{ij\}}^t \, , \{i,j\} \in A \tag{6l}$$

$$y_{\{ij\}}^t \leq Y_{\{ij\}}^t \qquad , \{i,j\} \in A \, , t \in T \tag{6m}$$

$x_{ij}^{skp} \in \{0,1\}$, $x_{ij}^{sl} \in \{0,1\}$, $z_i^{sr} \in \{0,1\}$, $y_{\{ij\}}^t$ inteiro não negativo

As restrições (6b), (6c) e (6d) asseguram que as variáveis x_{ij}^{skp} para cada VPC intra-região *k* definem um trajecto do seu nó origem para o seu nó destino. De notar que as restrições (6b) garantem que para cada VPC intra-região, existe um arco de saída na sua origem pelo menos para um valor de *p*. Da mesma forma, as restrições (6c) garantem que para cada VPC intra-região, existe um arco de entrada no seu destino pelo menos para um valor de *p*. As restrições (6d) garantem que nos outros nós da rede, para cada variável x_{ij}^{skp} que entra no nó e que esteja a 1, tem de haver uma outra variável x_{ij}^{skp} a 1 que saia do nó com o mesmo índice *p* da anterior. Assim, as restrições (6b), (6c) e (6d) no seu conjunto garantem que a solução final das variáveis x_{ij}^{skp} inclui um trajecto da origem para o destino, para pelo menos um dos valores do índice *p*.

As restrições (6e), (6f) e (6g) asseguram que para cada VPC inter-região *l*, as variáveis x_{ij}^{sl} definem um trajecto de um dos nós da região origem para um dos nós da região destino. Os VPCs inter-região são VPCs com extremos indefinidos e este tipo de restrições foi já apresentado na secção 2.3.1. As restrições (6h) e (6i), em conjunto com as 3 restrições anteriores, garantem que para cada região todos os VPCs inter-região que aí começam (6h) ou terminam (6i), fazem-no num único nó da região que é o nó fronteira (dado pelas variáveis z_i^{sr}).

As restrições (6j) e (6k) asseguram a consistencia dos índices *p* das variáveis x_{ij}^{skp} em relação à escolha do nó fronteira. Relembre-se que as restrições (6b-d) garantem a existência de um trajecto pelo menos para um valor de *p* das variáveis x_{ij}^{skp}. As restrições (6j) garantem que o arco de saída do trajecto no seu nó origem deverá ser definido por uma variável x_{ij}^{skp} com o índice $p = 1$ se a origem for o nó fronteira da região. Assim as

restrições (6j) em conjunto com (6b-d) garantem que existe um trajecto para o VPC definido pelas variáveis x_{ij}^{skp} com índice $p = 1$ se o nó origem do VPC for o nó fronteira da região. De modo análogo, as restrições (6k) em conjunto com (6b-d) garantem que existe um trajecto para o VPC definido pelas variáveis x_{ij}^{skp} com índice $p = 2$ se o nó destino do VPC for o nó fronteira da região. Se nenhum destes nós for o nó fronteira, então as variáveis x_{ij}^{skp} serão zero para os índices $p = 1$ e $p = 2$ e, pelas restrições (6b-d), deverá existir um trajecto para o VPC definido pelas variáveis x_{ij}^{skp} com índice $p = 0$.

As restrições (6l) garantem que a capacidade das ligações físicas em cada aresta é suficiente para suportar todos os VPCs (tanto os VPCs inter-região como os VPCs intra-região) de todas as redes lógicas que por lá passam. Ao contrário do capítulo anterior, neste caso consideramos apenas serviços simétricos pelo que a capacidade dos VPCs é igual nos dois sentidos. Assim, não é preciso aplicar as restrições (6l) a ambos os sentidos de cada aresta $\{i,j\}$ mas apenas a um dos seus sentidos. Finalmente, as restrições (6m) garantem que o número máximo de ligações físicas não é ultrapassado na solução encontrada.

Consideramos também uma pequena variação ao modelo apresentado em que se exige que o nó fronteira, em cada região, seja igual para todos os serviços. Este requisito poderá ser imposto pelos operadores por razões operacionais. Esta variante é facilmente introduzida no modelo apenas por modificação das variáveis que representam os nós fronteira. Nesta variante, substitui-se as variáveis binárias z_i^{sr} pelas variáveis binárias z_i^r que indicam se o nó $i \in N_r$ é o nó fronteira da região ou não. Na prática, o índice s das variáveis usadas no modelo apresentado é eliminado e garante-se assim que o nó fronteira de cada região é comum a todos os serviços.

4.5 Modelo de Dimensionamento Multi-Horário

No modelo multi-horário, cada VPC pode ser reconfigurado nos instantes de tempo entre períodos horários. Assim, tanto as grandezas que representam as capacidades dos VPCs como as variáveis binárias de trajecto (associadas a cada VPC) e de nó fronteira são definidas com um índice adicional que identifica o período horário a que se referem. A determinação dos valores de b_l^{sh} e b_k^{sph} deve ser feita para todos os períodos horários usando em cada um o procedimento descrito na secção 4.3. O modelo de programação

inteira que a seguir se apresenta é a extensão do modelo apresentado na secção anterior para o caso de redes com encaminhamento multi-horário:

N Conjunto de nós da rede

R Conjunto de regiões r da rede

N_r Conjunto dos nós de rede pertencentes à região $r \in R \left(N = \bigcup_r N_r \right)$

A Conjunto de arestas $\{i,j\}$ da rede: (i,j) representa a aresta $\{i,j\}$ no sentido de i para j e (j,i) representa o sentido inverso

T Conjunto dos tipos t de ligações físicas possíveis de implementar entre os nós de rede

α_t Largura de banda de uma ligação física do tipo t

$Y_{\{ij\}}^t$ Número máximo de ligações físicas do tipo t entre os nós i e j

$C_{\{ij\}}^t$ Custo associado à utilização de uma ligação do tipo t entre os nós i e j

S Conjunto de serviços s suportados pela rede

H Conjunto de períodos horários h

K_s Conjunto de VPCs intra-região k da rede lógica de suporte ao serviço s; cada VPC k é caracterizado por:

- o_k^s Nó origem
- d_k^s Nó destino
- r_k^s Região à qual pertence ($r_k^s \in R$)
- b_k^{sph} Para p = 0, capacidade no período horário h quando o VPC é interior; para p = 1, capacidade no período horário h quando o_k^s é o nó fronteira de r_k^s; para p = 2, capacidade no período horário h quando d_k^s é o nó fronteira de r_k^s

L_s Conjunto de VPCs inter-região l da rede lógica de suporte ao serviço s; cada VPC l é caracterizado por:

- o_l^s Região origem ($o_l^s \in R$)
- d_l^s Região destino ($d_l^s \in R$)
- b_l^{sh} Capacidade no período horário h

$y_{\{ij\}}^t$ Variável inteira que determina o número de ligações físicas do tipo t entre os nós i e j

x_{ij}^{slh} Variável binária de trajecto: o valor 1 significa que o VPC inter-região l de suporte ao serviço s passa no arco (i,j) no período horário h

x_{ij}^{skph} Variável binária de trajecto: para $p = 0$, o valor 1 significa que o VPC intra-região k no período horário h é interior e passa no arco (i,j); para $p = 1$, o valor 1 significa que no período horário h, o_k^s é o nó fronteira de r_k^s e o VPC passa no arco (i,j); para $p = 2$, o valor 1 significa que no período horário h, d_k^s é o nó fronteira de r_k^s e o VPC passa no arco (i,j)

z_i^{srh} Variável binária: o valor 1 significa que o nó $i \in N_r$ é o nó fronteira da região $r \in R$ na rede lógica de suporte ao serviço s no período horário h

Minimizar $$\sum_{\{i,j\}\in A}\sum_{t\in T} C_{\{ij\}}^t y_{\{ij\}}^t \tag{7a}$$

Sujeito a:

$$\sum_{j\in N:\{i,j\}\in A}\sum_{p=0}^{2}\left(x_{ij}^{skph} - x_{ji}^{skph}\right)=1 \quad , s \in S , k \in K_s , h \in H , i = o_k^s \tag{7b}$$

$$\sum_{j\in N:\{i,j\}\in A}\sum_{p=0}^{2}\left(x_{ji}^{skph} - x_{ij}^{skph}\right)=1 \quad , s \in S , k \in K_s , h \in H , i = d_k^s \tag{7c}$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{ij}^{skph} - x_{ji}^{skph}\right)=0 \quad , s \in S , k \in K_s , h \in H , p , i \in N \setminus \{ o_k^s , d_k^s \} \tag{7d}$$

$$\sum_{i\in N_r}\sum_{j\in N:\{i,j\}\in A}\left(x_{ij}^{slh} - x_{ji}^{slh}\right)=1 \quad , s \in S , l \in L_s , h \in H , r = o_l^s \tag{7e}$$

$$\sum_{i\in N_r}\sum_{j\in N:\{i,j\}\in A}\left(x_{ji}^{slh} - x_{ij}^{slh}\right)=1 \quad , s \in S , l \in L_s , h \in H , r = d_l^s \tag{7f}$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{ij}^{slh} - x_{ji}^{slh}\right)=0 \quad , s \in S , l \in L_s , h \in H , i \in N \setminus \left\{N_{o_l^s} , N_{d_l^s}\right\} \tag{7g}$$

$$z_i^{srh} = \sum_{j\in N:\{i,j\}\in A}\left(x_{ij}^{slh} - x_{ji}^{slh}\right) \quad , s \in S , r \in R , i \in N_r , h \in H , l \in L_s : o_l^s = r \tag{7h}$$

$$z_i^{srh} = \sum_{j\in N:\{i,j\}\in A}\left(x_{ji}^{slh} - x_{ij}^{slh}\right) \quad , s \in S , r \in R , i \in N_r , h \in H , l \in L_s : d_l^s = r \tag{7i}$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{ij}^{skph} - x_{ji}^{skph}\right)= z_i^{sr_k h} \quad , s \in S , k \in K_s , h \in H , i = o_k^s , p = 1 \tag{7j}$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{ji}^{skph} - x_{ij}^{skph}\right)= z_i^{sr_k h} \quad , s \in S , k \in K_s , h \in H , i = d_k^s , p = 2 \tag{7k}$$

$$\sum_{s\in S}\sum_{k\in K_s}\sum_{p=0}^{2} b_k^{sph}\left(x_{ij}^{skph}+x_{ji}^{skph}\right)+\sum_{s\in S}\sum_{l\in L_s} b_l^{sh}\left(x_{ij}^{slh}+x_{ji}^{slh}\right)\le\sum_{t\in T}\alpha_t y_{\{ij\}}^t \quad ,\{i,j\}\in A\,,\,h\in H \tag{7l}$$

$$y_{\{ij\}}^t \le Y_{\{ij\}}^t \qquad ,\{i,j\}\in A\,,\,t\in T \tag{7m}$$

$x_{ij}^{skph}\in\{0,1\}$, $x_{ij}^{slh}\in\{0,1\}$, $z_i^{srh}\in\{0,1\}$, $y_{\{ij\}}^t$ inteiro não negativo

A função de cada uma das restrições é semelhante à versão uni-horário deste modelo. As restrições (7b-k) são as restrições (6b-k) aplicadas a cada um dos períodos horários *h*. As restrições (7l) garantem que em cada período horário, a capacidade das ligações físicas para cada aresta é suficiente para suportar a soma de todas as capacidades dos VPCs desse período horário que por lá passam. O modelo apresentado permite que os nós fronteira sejam diferentes para diferentes serviços e para diferentes períodos horários. Facilmente se altera o modelo para garantir que os nós fronteira de cada região não variam de serviço para serviço (eliminando o índice *s* às variáveis z_i^{srh}) ou entre períodos horários (eliminando o índice *h* às variáveis z_i^{srh}) ou ambas as coisas.

4.6 Resultados Computacionais

Os resultados computacionais apresentados aplicam-se ao dimensionamento de uma rede com 10 nós e 16 arestas com a topologia apresentada na Figura 4-4 em que se considerou a possibilidade de utilizar ligações físicas PDH E3 de 34 Mbps entre os nós da rede e sem restrições no número máximo de ligações. Os custos $C_{\{ij\}}^t$ foram calculados atribuindo um custo de comutação de 3 para todos os nós e considerando ligações dedicadas com um custo de 0.2 por unidade de comprimento (os valores de comprimento estão apresentados na Figura 4-4).

Consideramos 2 serviços, cada um com 9 nós de tráfego. A largura de banda por ligação é de 128 Kbps para um serviço e 64 Kbps para o outro. Consideramos a probabilidade máxima de bloqueio de 1% para ambos os serviços. Para o modelo com redes lógicas hierarquizadas a dois níveis, dividimos a rede em 3 regiões. Consideramos a região 1 composta pelos nós 1, 2 e 3, a região 2 composta pelos nós 4, 5, 6 e 7 e a região 3 composta pelos restantes nós.

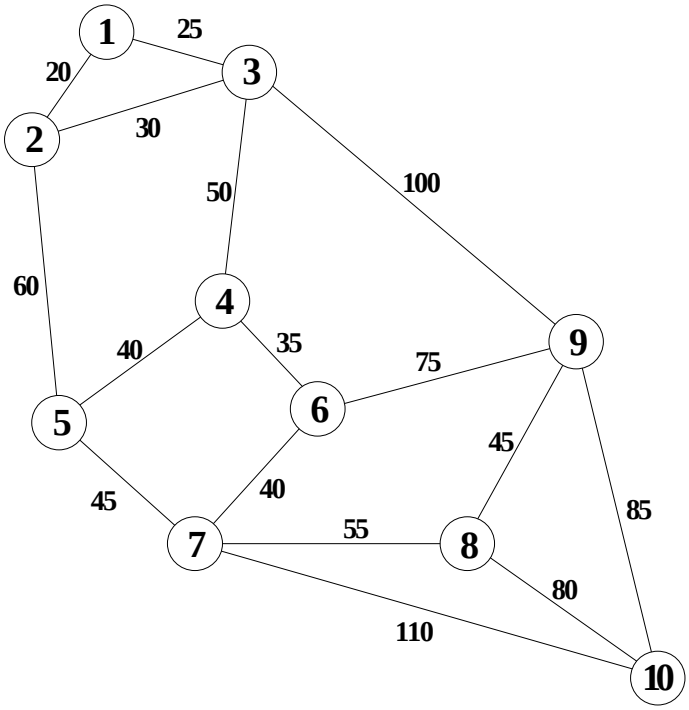


Figura 4-4 : Rede de teste

Foram gerados 10 problemas de dimensionamento. Em cada problema, foram seleccionados aleatoriamente 9 nós de tráfego para cada serviço e foram escolhidos aleatoriamente valores de tráfego para todos os pares de nós de serviço. Os valores de tráfego foram gerados com uma função densidade de probabilidade uniforme entre 1 e 10 Erlangs para os dois serviços.

A Tabela 4-1 apresenta os custos das soluções óptimas para os modelos uni-horário na configuração de redes lógicas com VPCs extremo-a-extremo e na configuração de redes lógicas hierarquizadas a dois níveis (forçando o mesmo nó fronteira e permitindo nós fronteira diferentes para cada serviço). Para o caso dos VPCs extremo-a-extremo, foi usada a formulação por modelização explícita de VPCs apresentada no capítulo 3. A tabela apresenta também os ganhos que se obtêm com os modelos hierarquizados a dois níveis para cada problema. A última linha apresenta os ganhos médios obtidos.

Os resultados computacionais obtidos mostram que os modelos hierarquizados a dois níveis permitem a implementação de redes físicas de custo menor do que as redes dimensionadas com base em redes lógicas com VPCs extremo-a-extremo. Os ganhos médios obtidos para os problemas considerados são de 12.1% quando se permitiu que os nós fronteira fossem diferentes para os dois serviços e 11.4% quando se impôs que fossem iguais. Estes resultados são significativos dado que foram obtidos para uma rede

relativamente pequena em que os ganhos de partilha de recursos são menores e as aproximações propostas são menos desprezáveis. De qualquer modo, os resultados mostram que, para os casos considerados, os ganhos de partilha de recursos das redes lógicas hierárquicas ultrapassa claramente o efeito da modularidade das ligações físicas.

	VPCs Extremo-a-Extremo	**Hierárquico Nós Front. Diferentes**	**Hierárquico Nós Front. Iguais**
1	183	176 (3.8%)	176 (3.8%)
2	204	176 (13.7%)	189 (7.4%)
3	214	192 (10.3%)	192 (10.3%)
4	196	176 (10.2%)	177 (9.7%)
5	195	178 (8.7%)	178 (8.7%)
6	216	178 (17.6%)	178 (17.6%)
7	206	164 (20.4%)	164 (20.4%)
8	204	176 (13.7%)	176 (13.7%)
9	210	191 (9.0%)	191 (9.0%)
10	206	178 (13.6%)	178 (13.6%)
Ganho médio:		12.1%	11.4%

Tabela 4-1 : Custo das soluções óptimas para os 10 problemas gerados

Os resultados mostram também que permitir nós fronteira diferentes para cada serviço conduz a soluções de custo menor ou igual ao custo que se obtém quando se impõe o mesmo nó fronteira. Este resultado é coerente com os modelos apresentados pois uma solução possível no modelo que impõe nós fronteira iguais é também solução no outro modelo. Embora na maioria dos casos as soluções encontradas tenham custos iguais, existe um problema (problema 2) em que o ganho é substancial quando se permite diferentes nós fronteira por serviço.

Os resultados foram obtidos através da ferramenta CPLEX usando o algoritmo de *branch-and-bound* a correr num PC a 400 MHz com 128 Mbytes de memória RAM. O cálculo prévio dos valores de capacidade foi da ordem dos segundos. O tempo máximo de cálculo para a resolução dos problemas apresentados foi de 4 horas embora a maioria dos problemas tenha sido resolvido em menos de 1 hora. O cálculo dos modelos hierarquizados a dois níveis foi em média três vezes mais rápido que o cálculo dos modelos com VPCs

extremo-a-extremo. Este facto indicia que os novos modelos são computationalmente mais eficazes.

Os tempos de cálculo dos problemas resolvidos significam que estes exemplos estão sensivelmente no limite do tamanho máximo dos problemas de dimensionamento que é possível calcular até à optimalidade com um algoritmo de *branch-and-bound*. Por esta razão, os resultados computacionais não incluem exemplos para cenários multi-horário nem redes com mais do que um tipo de ligações físicas. Foram realizados cálculos computacionais com redes de menor dimensão e com valores diferentes tanto para o número de serviços como para os valores médios de tráfego. Em muitos destes casos, as soluções com VPCs extremo-a-extremo foram melhores do que as soluções hierárquicas. Nestes casos (i) o número de classes de tráfego é reduzido e, por conseguinte, os ganhos de partinha de recursos são pequenos e (ii) as aproximações consideradas não são desprezáveis. Assim, o efeito da modularidade das ligações físicas torna-se determinante e justifica que as soluções com VPCs extremo-a-extremo sejam melhores.

Como conclusão, os resultados computacionais aqui apresentados não devem ser interpretados como uma prova de que as redes lógicas hierarquizadas a dois níveis conduzem sempre a melhores resultados de dimensionamento. Estes resultados apenas mostram que mesmo para redes pequenas é possível obter ganhos significativos com os modelos hierárquicos e indiciam que, para redes maiores, os ganhos podem ser bastante superiores quando comparados com soluções baseadas em VPCs extremo-a-extremo pois há maiores ganhos de partilha de recursos e as aproximações tornam-se menos significativas.

4.7 Conclusões

Neste capítulo, foi proposta uma arquitectura hierarquizada a 2 níveis para a topologia das redes lógicas de suporte a serviços conversacionais. Esta arquitectura, para além de permitir explorar ganhos de partilha de recursos, reduz significativamente o número de VPCs que é necessário configurar na rede e limita o número máximo de VPCs que cada ligação atravessa a um máximo de três (um factor determinante para serviços comutados em que o tempo de estabelecimento de ligação pode tornar-se crítico) sendo assim uma alternativa credível às redes baseadas em VPCs extremo-a-extremo.

Foram propostos dois modelos para o dimensionamento das redes com base nesta arquitectura: um modelo uni-horário e um modelo multi-horário. Nesta arquitectura não é possível separar o cálculo das capacidades dos VPCs do dimensionamento da rede. Foi no entanto mostrado como esta arquitectura permite, com pequenas aproximações, combinar de uma forma linear o cálculo das capacidades dos VPCs e o dimensionamento da rede física.

Os modelos de dimensionamento desenvolvidos permitiram calcular as soluções óptimas para redes até 10 nós. Os resultados computacionais obtidos mostram que a adopção de uma topologia hierarquizada a dois níveis pode conduzir a redes de custo significativamente menor quando comparadas com a tradicional configuração por VPCs extremo-a-extremo, mesmo para redes pequenas em que os ganhos de partilha de recursos são menores.

CAPÍTULO 5

REDES LÓGICAS EM ÁRVORE

Para além de não explorar ganhos de partilha de recursos, a utilização de redes lógicas baseadas em VPCs extremo-a-extremo no suporte de serviços ponto-multiponto tem ainda a desvantagem de não explorar a capacidade de replicação de células que os comutadores ATM permitem. Numa rede lógica com VPCs extremo-a-extremo, uma ligação ponto-multiponto é estabelecida pelos VPCs que terão necessariamente de existir entre o nó origem da ligação e cada um dos seus nós destinos (secção 3.1.3). Neste caso, o nó origem é o responsável por replicar as células recebidas pelo terminal chamador para todos os nós destino. Dependendo da posição na rede dos nós envolvidos na ligação estabelecida, esta forma de gestão de recursos pode ser pouco eficiente. Por exemplo, no caso em que os nós destino de uma ligação ponto-multiponto estão todos próximos uns dos outros e afastados do nó origem, esta arquitectura faz com que várias réplicas do mesmo fluxo de células ATM atravessem a rede da origem para a zona dos nós destinos. Neste caso, uma utilização mais eficiente dos recursos seria enviar um único fluxo de células para um dos nós destino e esse nó replicar as células para os outros nós.

Neste capítulo é proposta uma alternativa à arquitectura com VPCs extremo-a-extremo em que a rede lógica de suporte é constituída por um conjunto de VPCs ponto-a-ponto com uma arquitectura em árvore. Cada serviço divide os nós da rede ATM em dois

conjuntos: os nós de tráfego (nós em que o serviço é oferecido a clientes) e nós de trânsito (nós em que o serviço não é oferecido a clientes). Os VPCs que formam a rede lógica deverão interligar todos os nós de tráfego do serviço. O problema abordado nesta capítulo consiste então em dimensionar redes ATM para suporte a serviços ponto-multiponto suportados por redes lógicas com uma arquitectura em árvore. A determinação da árvore que define cada rede lógica é um dos resultados do problema de dimensionamento a formular e, no caso de redes com configuração dinâmica de VPCs (modelo multi-horário), permite-se que as redes sejam alteradas de período para período horário.

A arquitectura proposta neste capítulo não permite separar o cálculo das capacidades dos VPCs do dimensionamento de rede física como no caso dos VPCs extremo-a-extremo. Também não é possível combinar o cálculo das capacidades dos VPCs com o dimensionamento da rede física como no caso das redes lógicas hierárquicas porque não se sabe à partida entre que nós de tráfego os VPCs serão estabelecidos nem tão pouco que classes de tráfego atravessam cada VPC. No caso de redes lógicas em árvore, é necessário modelizar os trajectos das classes de tráfego através de variáveis de trajecto e é necessário determinar um conjunto de restrições adequadas que garantam que a capacidade dos VPCs cumpre com a qualidade de serviço pretendida para cada serviço (nomeadamente a probabilidade de bloqueio máxima das ligações lógicas). Este conjunto de restrições é então incorporado no modelo de dimensionamento da rede que simultaneamente calcula as capacidades e os trajectos dos VPCs e a rede física resultante. Neste capítulo são propostas duas formas lineares de formulação de restrições de probabilidade de bloqueio, ao nível da ligação lógica, na determinação das capacidades dos VPCs que permitem modelizar o problema de dimensionamento da rede de uma forma linear.

Tal como foi referido no capítulo anterior para as redes lógicas hierárquicas, o efeito da modularidade das ligações físicas é maior nas redes lógicas em árvore do que nas redes com VPCs extremo-a-extremo. De facto, estas duas arquitecturas constituem os dois extremos em relação ao número de VPCs que produzem e portanto, o efeito da modularidade poderá ser determinante na comparação dos resultados de dimensionamento.

Este capítulo está organizado da seguinte forma. A secção 5.1 descreve a abordagem ao problema de dimensionamento. Nesta secção é apresentado o método pelo qual é possível formular de uma forma linear as restrições de probabilidade de bloqueio. Como consequência deste método, são propostos dois tipos de restrições que servem de base aos

modelos de dimensionamento da rede apresentados na secção 5.2. Na secção 5.3 são apresentadas reformulações válidas dos modelos da secção anterior. A secção 5.4 apresenta os modelos adequados ao caso multi-horário. Finalmente, os resultados computacionais são apresentados na secção 5.5 em que é analisada a eficiência dos diferentes modelos propostos e são comparados os resultados do dimensionamento da rede entre o uso de redes lógicas com VPCs extremo-a-extremo e redes lógicas em árvore.

5.1 Abordagem ao Problema

Uma rede lógica em árvore é uma rede lógica constituída por um conjunto mínimo de VPCs que interliga sem ciclos os nós de tráfego do serviço suportado. Sendo J o conjunto de nós de tráfego do serviço suportado, qualquer conjunto de $|J| - 1$ VPCs que interligue todos os nós constitui uma rede lógica em árvore.

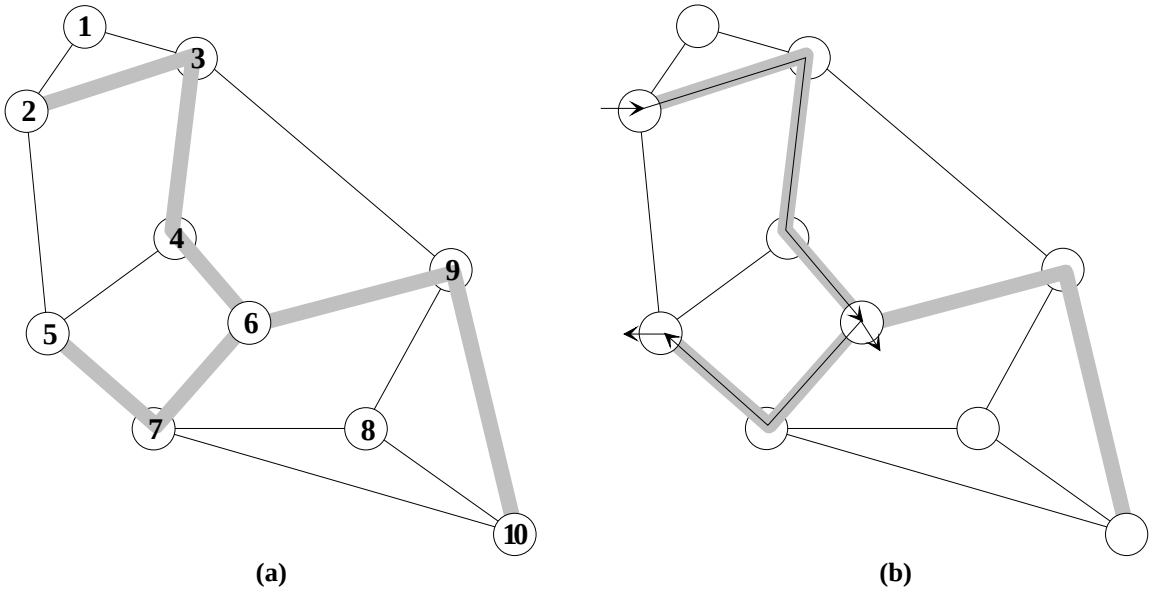


Figura 5-1 : Exemplo de uma rede lógica em árvore (a) e do estabelecimento de um ligação ponto-multiponto (b)

A Figura 5-1a ilustra uma rede lógica com uma arquitectura em árvore em que os nós de tráfego do serviço suportado são os nós 2, 5, 6 e 10. Esta rede é constituída por 3 VPCs ponto-a-ponto: um VPC entre os nós 2 e 6, um VPC entre os nós 5 e 6 e um VPC entre os nós 6 e 10. Dado que uma árvore de VPCs constitui uma arquitectura sem ciclos, ela define univocamente a árvore direccionada pela qual uma ligação ponto-multiponto é estabelecida. No exemplo apresentado na Figura 5-1b, uma ligação do nó 2 para os nós 5 e

6 é estabelecida pelo conjunto de VPCs que interliga os 3 nós envolvidos resultando na árvore direccionada definida pelos arcos (2,3), (3,4), (4,6), (6,7) e (7,5). Neste exemplo, o nó 6 é responsável pela cópia do fluxo vindo do nó origem para os 2 destinos da ligação.

No caso de redes lógicas em árvore, o cálculo das capacidades de cada VPC é dependente da topologia da rede lógica porque os extremos de cada VPC não são conhecidos à partida. Assim, o maior desafio na abordagem a este problema é conseguir incluir o cálculo da capacidade dos VPCs (dependente da qualidade de serviço necessária ao nível de ligação) na formulação do problema de dimensionamento da rede e, no entanto, conseguir que o modelo de optimização resultante seja linear. Este objectivo é conseguido em dois passos.

A capacidade de cada VPC depende da probabilidade de bloqueio do tráfego que por ele passa. Numa arquitectura em que as classes de tráfego atravessam mais do que um VPC, a probabilidade de bloqueio de um VPC é dependente da probabilidade de bloqueio que as classes de tráfego que o atravessam sofrem nos outros VPCs. Assim, a capacidade de um VPC é dependente das capacidades dos outros VPCs que constituem a rede lógica. O primeiro passo é aplicar o teorema do limite do produto (conhecido da literatura para tráfego ponto-a-ponto) ao caso de tráfego ponto-multiponto. Este teorema é uma aproximação conservativa e resulta num modelo em que a capacidade de cada VPC deixa de ser dependente da capacidade dos outros VPCs e é apenas determinada pela função ErlangB considerando as classes de tráfego que por ele atravessam.

O segundo passo é determinar um conjunto de equações lineares que representem a função ErlangB inversa. Definidos estes dois passos, é possível determinar modelos lineares para o problema de dimensionamento da rede.

5.1.1 Majorante da Probabilidade de Bloqueio

Recordemos o teorema do limite do produto aplicado a classes de tráfego ponto-a-ponto [Whitt85]. Para o podermos enunciar, introduzimos a seguinte notação que se refere a uma rede lógica de suporte a um serviço ponto-a-ponto simétrico:

K Conjunto dos VPCs que constituem a rede lógica de suporte

m_k Número de ligações simultâneas suportado pelo VPC k

F Conjunto das classes de tráfego f; cada classe de tráfego f é definida por:

ρ_f Intensidade de tráfego (em Erlangs)

R_f Conjunto dos VPCs *k* que definem o seu trajecto

B_f Probabilidade de bloqueio da classe de tráfego *f*

F_k Conjunto das classes de tráfego *f* que atravessam o VPC $k \rightarrow F_k = \{f : k \in R_f\}$

Com base nesta notação e recorrendo ao teorema do limite do produto [Whitt85], a probabilidade de bloqueio de cada classe de tráfego *f* é limitada superiormente por:

$$B_f \leq 1 - \prod_{k \in R_f} (1 - L_k) \tag{1}$$

em que

$$L_k = ER\left[\sum_{f \in F_k} \rho_f, m_k\right] \tag{2}$$

Para casos em que os valores de probabilidade de bloqueio B_f são pequenos e as classes de tráfego atravessam poucos VPCs ($|R_f|$ pequeno para todas as classes de tráfego *f*), o limite do produto é uma boa aproximação. Uma forma de interpretar este teorema é considerar que os valores L_k , dados por (2), são a probabilidade de bloqueio que as classes de tráfego sofrem ao atravessar o VPC *k* e que a probabilidade de bloqueio total de cada classe de tráfego, dada por (1), é um menos a probabilidade de não haver bloqueio em nenhum dos VPCs que atravessa.

Considere-se agora o caso de tráfego ponto-multiponto. Como as classes de tráfego ponto-multiponto são uni-direccionais (secção 2.4.2), temos de considerar os dois sentidos de cada VPC como dois VPCs distintos e o conjunto *K* representa o conjunto dos VPCs direccionados da rede lógica. Por exemplo, considere-se a rede lógica definida na Figura 5-1. Definindo o conjunto de VPCs pelo par ordenado (*i*,*j*) em que *i* é a origem e *j* é o destino, o conjunto *K* é dado por:

K = {(2,6),(6,2),(5,6),(6,5),(6,10),(10,6)}

As classes de tráfego ponto-multiponto são representadas por f_o^r em que *o* é o nó origem e *r* é o conjunto de nós destino. R_o^r é o conjunto dos VPCs *k* que definem a árvore de estabelecimento das ligações de f_o^r. Ainda no exemplo da rede lógica definida na Figura 5-1, se considerarmos um cenário de tráfego com as seguintes classes de tráfego:

$F = \{ f_2^{5,6}, f_5^{10}, f_6^{5,10}, f_{10}^{5,6}, f_{10}^{2,5,6} \}$

então, os conjuntos R_o^r são dados por:

$R_2^{5,6}$ = {(2,6),(6,5)} R_5^{10} = {(5,6),(6,10)}

$R_6^{5,10}$ ={(6,5),(6,10)} $R_{10}^{5,6}$ = {(10,6),(6,5)}

$R_{10}^{2,5,6}$ = {(10,6),(6,5),(6,2)}

e os conjuntos F_k são dados por:

$F_{(2,6)}$ = { $f_2^{5,6}$ } $F_{(6,2)}$ = { $f_{10}^{2,5,6}$ }

$F_{(5,6)}$ = { f_5^{10} } $F_{(6,5)}$ = { $f_2^{5,6}$, $f_6^{5,10}$, $f_{10}^{5,6}$, $f_{10}^{2,5,6}$ }

$F_{(6,10)}$ = { f_5^{10} , $f_6^{5,10}$ } $F_{(10,6)}$ = { $f_{10}^{5,6}$, $f_{10}^{2,5,6}$ }

Numa rede lógica de suporte a um serviço com encaminhamento fixo, a árvore de serviço do nó *o* é a árvore direccionada definida do nó *o* para todos os outros nós de tráfego do serviço pela qual a rede tenta estabelecer todas as chamadas ponto-multiponto submetidas em *o*. Representemos a árvore de serviço por R_o. Na rede lógica definida na Figura 5-1, existem quatro árvores de serviço R_o, tantos quantos o número de nós de tráfego do serviço. Os conjuntos R_o são dados por:

R_2 = {(2,6),(6,5),(6,10)} R_5 = {(5,6),(6,2),(6,10)}

R_6 ={(62),(6,5),(6,10)} R_{10} = {(10,6),(6,2),(6,5)}

Pela definição, qualquer conjunto R_o^r está incluído em R_o. Representemos também por R_o^d o sub-conjunto de R_o que define o trajecto do nó origem *o* para o nó destino *d*. Note-se que, conforme foi definido na secção 2.5, a probabilidade de bloqueio de um nó origem *o* para um nó destino *d* é a probabilidade de uma chamada ponto-multiponto de *o* para qualquer conjunto de nós destino que inclua *d* não ser estabelecida para *d*. Assim, é sobre o trajecto R_o^d que podemos agora definir a probabilidade de bloqueio do serviço.

De seguida, é resumida a notação apresentada (algumas grandezas são as mesmas que a notação apresentada para o tráfego ponto-a-ponto e são aqui repetidas por questões de clareza):

K Conjunto dos VPCs direccionados que constituem a rede lógica de suporte

m_k Variável inteira: número de ligações simultâneas suportado pelo VPC *k*

F Conjunto das classes de tráfego f_o^r em que *o* é o nó origem e *r* é o conjunto de nós destino; a classe de tráfego f_o^r é ainda definida por:

- ρ_o^r Intensidade de tráfego (em Erlangs)
- R_o^r Conjunto dos VPCs *k* que definem a sua árvore direccionada ($R_o^r \subset R_o$)

R_o Árvore de serviço do nó *o*

R_o^d Sub-conjunto de R_o que define o trajecto do nó origem *o* para o nó destino *d*

B_o^d Probabilidade de bloqueio do nó origem *o* para o nó destino *d*

F_k Classes de tráfego f_o^r que atravessam o VPC $k \rightarrow F_k = \left\{ f_o^r : k \in R_o^r \right\}$

Com base nesta notação e recorrendo novamente ao teorema do limite do produto, a probabilidade de bloqueio de cada origem *o* para cada destino *d* é limitada superiormente por:

$$B_o^d \leq 1 - \prod_{k \in R_o^d} (1 - L_k) \tag{3}$$

em que

$$L_k = ER\left[\sum_{f_o^r \in F_k} \rho_o^r, m_k \right] \tag{4}$$

As equações (3) e (4) definem a relação entre o tráfego ρ_o^r das classes submetidas à rede lógica, a probabilidade de bloqueio de cada par origem-destino B_o^d e as capacidades m_k dos VPCs que compõem a rede lógica.

O diâmetro de uma rede lógica em árvore é definido como o número máximo de VPCs que interliga qualquer par de nós de tráfego da rede lógica. O diâmetro de uma rede lógica em árvore está sempre compreendido entre 2 (uma árvore em estrela) e o número total de VPCs (uma árvore em linha). No exemplo da Figura 5-1, o diâmetro é 2 porque o trajecto na árvore que liga quaisquer par de nós de tráfego é composto no máximo por dois VPCs. Na prática, o diâmetro máximo da rede lógica está limitado pelo tempo máximo de estabelecimento de ligação que o operador pretende e é, por isso, um parâmetro de entrada do problema de dimensionamento. O tempo de estabelecimento de uma ligação ponto-multiponto é directamente proporcional ao número de VPCs entre o nó origem e o nó destino mais distante. Recorrendo novamente ao exemplo da Figura 5-1, considere-se o estabelecimento de uma ligação ponto-multiponto do nó 2 para os nós 6 e 5 (Figura 5-1b). O nó 2 recebe o pedido do terminal chamador e verifica se tem recursos no VPC (2,6) para suportar a ligação. Se tiver recursos, embora saiba desde logo que pode estabelecer a ligação para o nó 6, precisa de encaminhar o pedido para este nó pois só ele pode determinar se tem recursos no VPC (6,5) para estabelecer a ligação completa. Assim, o tempo de estabelecimento de chamada envolve 2 saltos que é igual ao número de VPCs que interliga o nó 2 ao nó destino mais distante, ou seja o nó 5. Por esta razão, o diâmetro máximo da árvore tem de ser um parâmetro a ter em conta na determinação da rede lógica

de suporte a qualquer serviço ponto-multiponto. Considerando este parâmetro dado por *M*, então:

$|R_o^d| \leq M$, para todos os conjunto R_o^d.

Então, se a qualidade de serviço ao nível da ligação lógica exigir uma probabilidade de bloqueio menor ou igual a Γ para todos os pares origem destino, a equação 3 pode ser observada impondo uma probabilidade de bloqueio para todos os VPCs menor ou igual do que um valor P_L dado por:

$$\Gamma = 1-(1-P_L)^M \Leftrightarrow P_L = 1-\sqrt[M]{1-\Gamma} \tag{5}$$

e, portanto, fazendo para todos os VPCs *k*:

$$ER\left[\sum_{f_o^r \in F_k} \rho_o^r, m_k\right] \leq P_L \tag{6}$$

Esta aproximação é tanto melhor quanto menor for o valor *M*. Se definirmos uma variável binária x_k^{or} cujo conteúdo, se for 1, indica que a classe de tráfego do nó *o* para o conjunto de nós *r* atravessa o VPC *k*, então podemos reescrever a equação (6) na seguinte forma:

$$ER\left[\sum_o \sum_r \left(\rho_o^r \cdot x_k^{or}\right), m_k\right] \leq P_L \text{ , para todos os } k \in K \tag{7}$$

As equações (7) definem, para cada VPC, uma relação não linear (dada pela fórmula de ErlangB) entre a soma das intensidades de tráfego que atravessam o VPC *k* e a sua capacidade m_k em número de ligações simultâneas. As equações (7) garantem que a probabilidade de bloqueio máxima do serviço suportado é cumprida e eliminam a dependência entre o cálculo da capacidade de cada VPC e a probabilidade de bloqueio que as classes de tráfego que o atravessam sofrem nos outros VPCs. O passo seguinte é determinar um conjunto de equações lineares que garantam que as equações (7) sejam observadas na solução final.

5.1.2 Modelos Lineares da Função ErlangB Inversa

Considerando a notação habitual da função ErlangB P = ER[ρ,m], representa-se a sua função inversa relativamente a m por m = ER^{-1}[ρ,P]. Esta função não tem uma fórmula fechada e é apenas possível ser calculada de uma forma algorítmica. Existem, no entanto,

algoritmos eficientes para a calcular [Jagerman84]. A Figura 5-2 e a Figura 5-3 ilustram a função $ER^{-1}[\rho,P]$. As figuras apresentam o número de ligações em função do tráfego para três valores diferentes de probabilidade de bloqueio. A Figura 5-2 é uma ampliação da Figura 5-3 para os valores de intensidade de tráfego até 6 Erlangs.

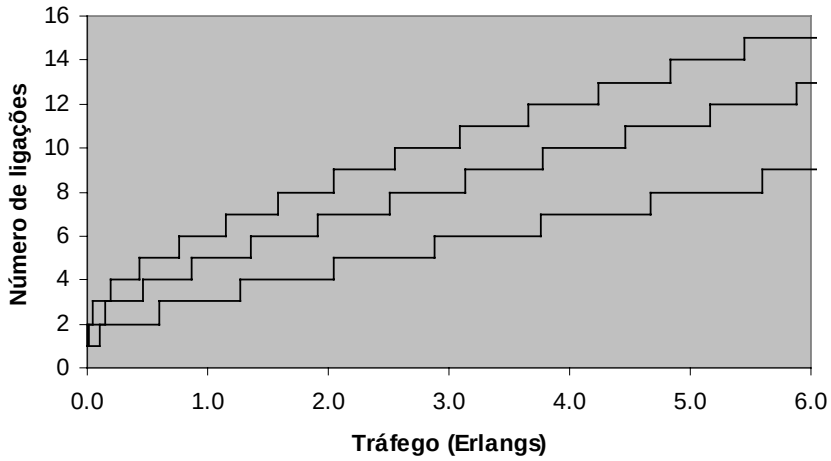


Figura 5-2 : $ER^{-1}[\rho,P]$ para tráfego até 6.0 Erlangs

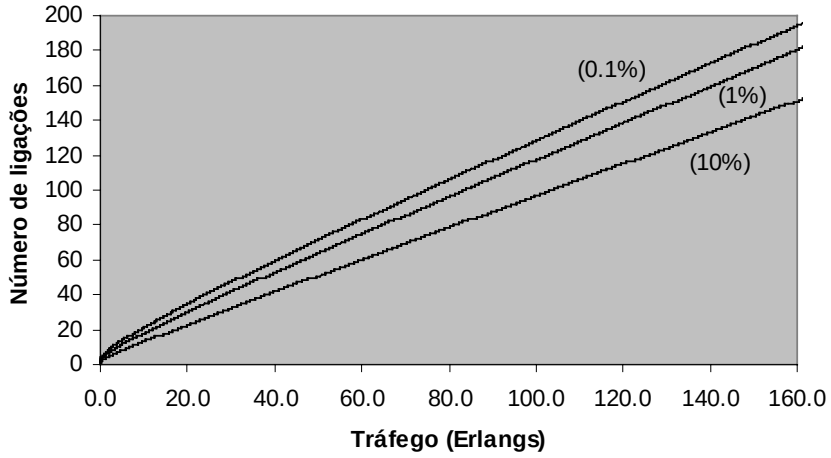


Figura 5-3 : $ER^{-1}[\rho,P]$ para tráfego até 160.0 Erlangs

Conforme ilustrado pelas figuras apresentadas, para o mesmo valor de intensidade de tráfego, um número de ligações maior resulta numa probabilidade de bloqueio menor. Assim, para que as desigualdades definidas pelas equações (7) sejam observadas, é preciso que:

$$m_k \geq ER^{-1}\left[\sum_o \sum_r \left(\rho_o^r \cdot x_k^{or}\right), P_L\right] \quad \text{, para todos os } k \in K \tag{8}$$

Esta relação pode ser descrita de uma forma exacta ou de uma forma aproximada. A forma exacta recorre exclusivamente a variáveis binárias e designamos o modelo resultante por modelo binário. A forma aproximada recorre a um conjunto de váriáveis que contempla variáveis binárias e variáveis reais e designamos o modelo resultante por modelo real. De seguida, consideramos cada modelo separadamente.

5.1.3 Modelo Binário da Função ErlangB Inversa

Para um valor fixo de probabilidade de bloqueio, a função $ER^{-1}[\rho,P]$ (Figura 5-2) é não decrescente e em forma de degrau. Assim, considere-se Ω_n na função $ER^{-1}[\rho,P]$, o valor de tráfego no qual a função tem o degrau de n para n+1 no número de ligações. Considerem-se também as variáveis binárias u_k^n cujo conteúdo se for 1, indica que o VPC tem capacidade para n ligações em simultâneo. Com base em (8) e introduzindo Ω_n e u_k^n é possível enunciar a relação entre a capacidade de cada VPC e as classes de tráfego que o atravessam da seguinte forma:

$$\sum_o \sum_r \rho_o^r \cdot x_k^{or} \le \sum_n \Omega_n \cdot u_k^n \qquad \text{, para todos os } k \in K \tag{9}$$

$$\sum_n u_k^n \le 1 \qquad \text{, para todos os } k \in K \tag{10}$$

em que os valores Ω_n são calculados para uma probabilidade de bloqueio P_L dada por (5). A equação (10) garante que para cada VPC k no máximo uma das variáveis u_k^n é um e a equação (9) garante que, para a probabilidade de bloqueio P_L, o número de ligações do VPC é suficiente para suportar o tráfego que o atravessa. Note-se que como foi assumido que todos os VPCs têm probabilidade de bloqueio máxima igual, os valores Ω_n são iguais para todos os VPCs de uma rede lógica. Este modelo gera um número significativo de variáveis binárias pelo que é importante estimar um limite inferior e um limite superior para os valores de n. Designando estes valores por $\underline{n}$ (limite inferior) e $\overline{n}$ (limite superior), as equações anteriores podem ser reescritas na seguinte forma:

$$\sum_o \sum_r \rho_o^r \cdot x_k^{or} \le \sum_{n=\underline{n}}^{\overline{n}} \Omega_n \cdot u_k^n \qquad \text{, para todos os } k \in K \tag{11}$$

$$\sum_{n=\underline{n}}^{\overline{n}} u_k^n \le 1 \qquad \text{, para todos os } k \in K \tag{12}$$

As equações (11) e (12) podem agora ser usadas num modelo linear de forma a garantir que com base nas variáveis *x* que definem a árvore (ou trajecto) de cada classe de tráfego, a capacidade (em número de ligações) dos VPCs que atravessam (definida pelas variáveis binárias *u*) é suficiente para cumprir com a probabilidade de bloqueio desejada.

Resta agora descrever de que forma são calculados os valores de $\underline{n}$ (limite inferior) e $\overline{n}$ (limite superior) a usar nas equações introduzidas.

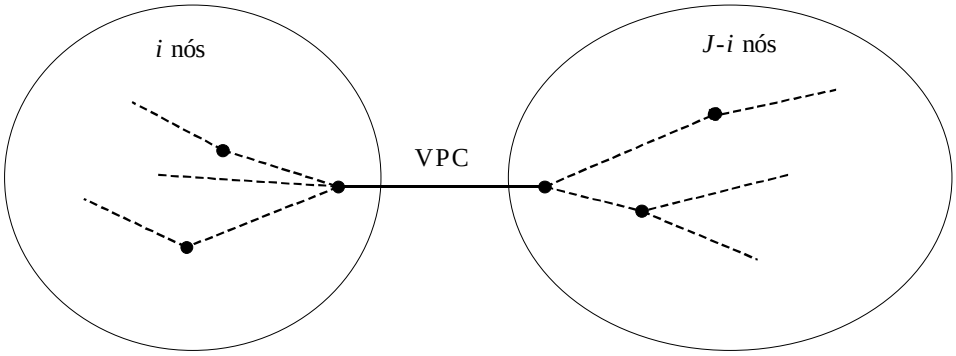


Figura 5-4 : Cada VPC divide uma rede lógica em duas partes

Relembre-se que nos cenários de tráfego de interesse (secção 2.4.2), considera-se que numa rede de *N* nós, existem *J* nós de tráfego entre os quais existem classes de tráfego de todos os *J* nós para todos os conjuntos de no máximo *P* nós destinos. Assim, o número de classes de tráfego de um dado nó de tráfego para *p* destinos ($1 \le p \le P$) é dado pelo número de combinações de *p* nós nos restantes $J - 1$ nós:

$$\binom{J-1}{p}$$

Um VPC de uma rede lógica em árvore divide o conjunto de *J* nós em duas partes (conforme ilustrado na Figura 5-4). O número de classes de tráfego para *p* destinos com origem do lado esquerdo do VPC da Figura 5-4 é igual a:

$$i \cdot \binom{J-1}{p}$$

e destas classes de tráfego, as que não atravessam o VPC são aquelas em que todos os destinos pertencem ao mesmo lado da origem. Assim, o número de classes de tráfego para *p* destinos que atravessa o VPC (da esquerda para a direita da Figura 5-4) é dado por:

$$i \cdot \binom{J-1}{p} - i \cdot \binom{i-1}{p} \qquad \text{, se } i > p \tag{13a}$$

$$i \cdot \binom{J-1}{p} \qquad \text{, se } i \le p \tag{13b}$$

em que quando $i \le p$, todas as classes de tráfego cuja origem está no lado esquerdo da Figura 5-4 atravessam o VPC.

Fazendo variar i de 1 até $J - 1$ podemos registar o menor e o maior de todos os valores da expressão (13). Estes são os valores mínimo e máximo de classes de tráfego com p destinos que atravessam qualquer VPC de uma rede lógica em árvore. Designamos estes valores por $\underline{a}$ e $\overline{a}$. Somando os $\underline{a}$ menores valores de intensidade de tráfego de todas as classes de tráfego para p destinos, obtém-se um limite inferior para o tráfego que passa em cada VPC para p destinos. Se repetirmos este procedimento para cada valor de p e somarmos todos os valores, obtemos um limite inferior para o tráfego que passa por qualquer VPC. De forma análoga é possível obter um limite superior: para cada valor de p, somam-se os $\overline{a}$ maiores valores de intensidade de tráfego de todas as classes de tráfego para p destinos, de seguida, somam-se todos estes valores e obtém-se um limite superior para o tráfego que passa por qualquer VPC.

Com cada um destes valores limite, é possível através de (8) calcular os valores de $\underline{n}$ (limite inferior) e $\overline{n}$ (limite superior) para o número de ligações dos VPCs da rede lógica. O método de cálculo de $\underline{n}$ e $\overline{n}$ proposto é genericamente descrito pelo seguinte algoritmo:

Fazer TráfegoMínimo = 0

Fazer TráfegoMáximo = 0

Para p de 1 até P executar:

- Percorrer i de 1 até $J - 1$ para obter o valor mínimo ($\underline{a}$) e máximo ($\overline{a}$) da expressão (13)
- Fazer TráfegoMínimo = TráfegoMínimo + as $\underline{a}$ menores intensidades de tráfego de todas as classes de tráfego para p destinos
- Fazer TráfegoMáximo = TráfegoMáximo + as $\overline{a}$ maiores intensidades de tráfego de todas as classes de tráfego para p destinos

Calcular $\underline{n}$ como o menor inteiro maior do que ER^{-1}[TráfegoMínimo,P_L]

Calcular $\overline{n}$ como o menor inteiro maior do que ER^{-1}[TráfegoMáximo,P_L]

5.1.4 Modelo Real da Função ErlangB Inversa

Devido ao facto dos modelos binários introduzirem uma quantidade apreciável de variáveis inteiras (que em termos de modelos de programação tornam a sua resolução bastante mais difícil), consideramos também a utilização de uma semi-recta como aproximação à função ErlangB inversa. Conforme ilustrado na Figura 5-3, a partir de valores razoáveis de intensidade de tráfego, esta função pode ser aproximada a uma recta com um erro pequeno. Assim, o objectivo é reescrever as equações (11) e (12) na seguinte forma:

$$\sum_o \sum_r \rho_o^r \cdot x_k^{or} \leq \Phi_{\underline{n}} \cdot u_k + \Phi_\Delta \cdot u_k^\Delta \quad \text{, para todos os } k \in K \tag{14}$$

$$u_k \geq 0.001 \cdot u_k^\Delta \quad \text{, para todos os } k \in K \tag{15}$$

em que a variável u_k é binária e a variável u_k^Δ é real. Na equação (14), o somatório das variáveis binárias multiplicadas pelos respectivos coeficientes da equação (11) é substituído pela expressão de uma semi-recta em que $\Phi_{\underline{n}}$ é o valor inicial e Φ_Δ é o valor do declive. A variável binária u_k representa, se estiver a um, que o VPC *k* tem pelo menos uma capacidade de $\underline{n}$ ligações e a variável real u_k^Δ determina a capacidade adicional do VPC *k*. Este modelo reduz significativamente o número de variáveis introduzidas no modelo tornando o modelo de programação resultante de menor complexidade. A equação (15) garante que, para que a variável real u_k^Δ possa ser maior que zero, é necessário que a variável binária u_k seja um. O coeficiente da variável u_k^Δ usado na equação (15) deve ser suficientemente pequeno para que o limite máximo desta variável não seja imposto por esta equação. Dado que u_k é no máximo um, a escolha do coeficiente 0.001 implica que a variável u_k^Δ é no máximo 1000. Este valor é suficientemente grande para os problemas de dimensionamento abordados nesta tese.

Consideremos agora a determinação dos valores de $\Phi_{\underline{n}}$ e Φ_Δ. Para que a aproximação seja conservativa, é suficiente que a semi-recta definida por estes valores inclua os vértices de dois degraus consecutivos da função original. A Figura 5-5 ilustra esta afirmação. Na Figura 5-5a, os vértices utilizados para definir a semi-recta são os dois primeiros vértices (correspondentes aos degraus de $\underline{n}$ e $\underline{n}$+1) e a Figura 5-5b apresenta a semi-recta definida pelos vértices correspondentes aos degraus de $\underline{n}$+5 e $\underline{n}$+6.

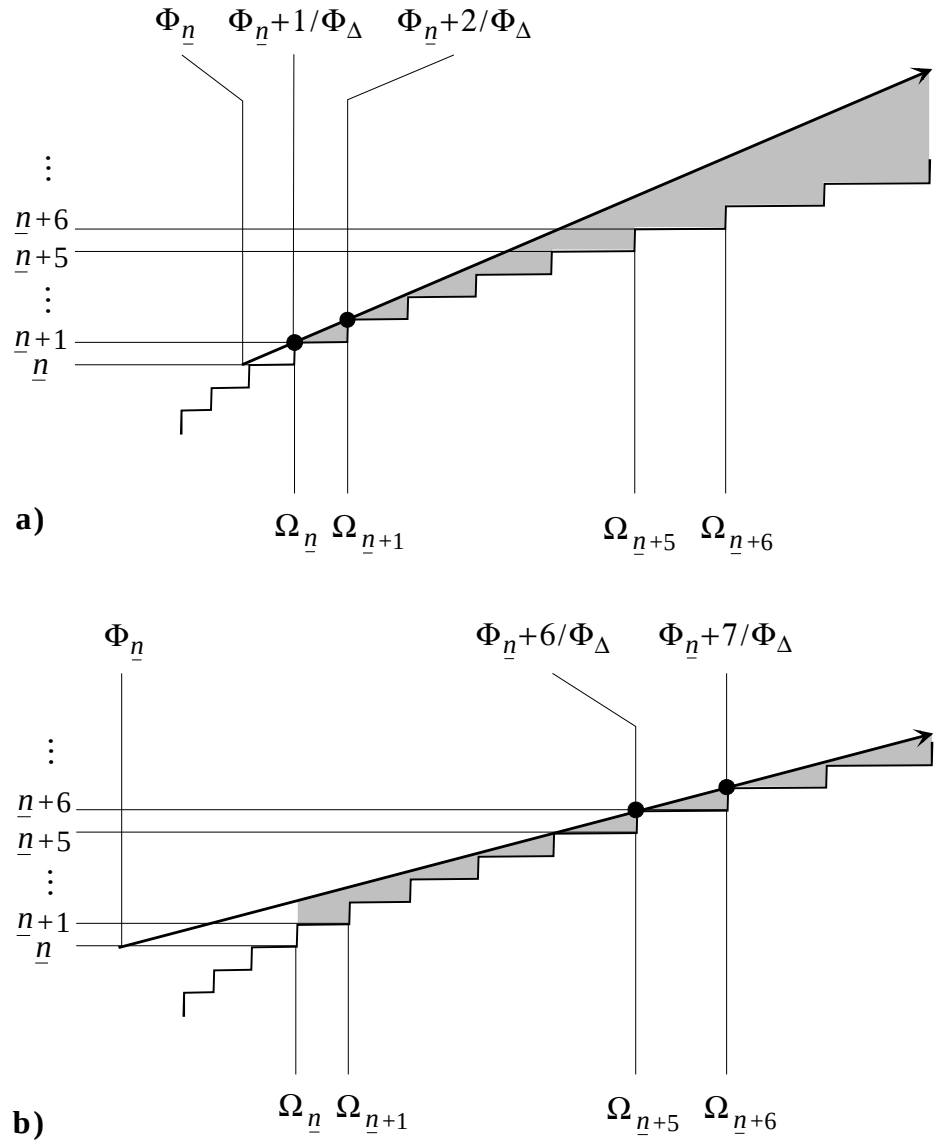


Figura 5-5 : Ilustração da aproximação de $ER^{-1}[\rho,P]$ por uma semi-recta

Para que a semi-recta definida pelos valores de $\Phi_{\underline{n}}$ e Φ_Δ inclua o vértice *i*, é necessário que a seguinte equação seja cumprida:

$$\underline{n}+\Phi_\Delta\left(\Omega_i-\Phi_{\underline{n}}\right)=i+1 \quad \Leftrightarrow \quad \Phi_{\underline{n}}+\left(i-\underline{n}+1\right)/\Phi_\Delta=\Omega_i$$

Assim, os valores de $\Phi_{\underline{n}}$ e Φ_Δ da semi-recta que passa pelos vértices *i* e *i*+1 (com $\underline{n} \le i < \overline{n}-1$) são dados pelo sistema de duas equações a duas incognitas seguinte:

$$\begin{cases} \Phi_{\underline{n}}+\left(i-\underline{n}+1\right)/\Phi_\Delta=\Omega_i \\ \Phi_{\underline{n}}+(i-\underline{n}+2)/\Phi_\Delta=\Omega_{i+1} \end{cases} \tag{16}$$

A distância entre a função original e a recta escolhida dá, para cada valor de tráfego, o sobre-dimensionamento na capacidade do VPC que esta aproximação introduz (Figura

5-6). Verifica-se facilmente pelos 2 exemplos apresentados na Figura 5-5, que para cada valor de tráfego, o sobre-dimensionamento introduzido é diferente para diferentes semi-rectas escolhidas.

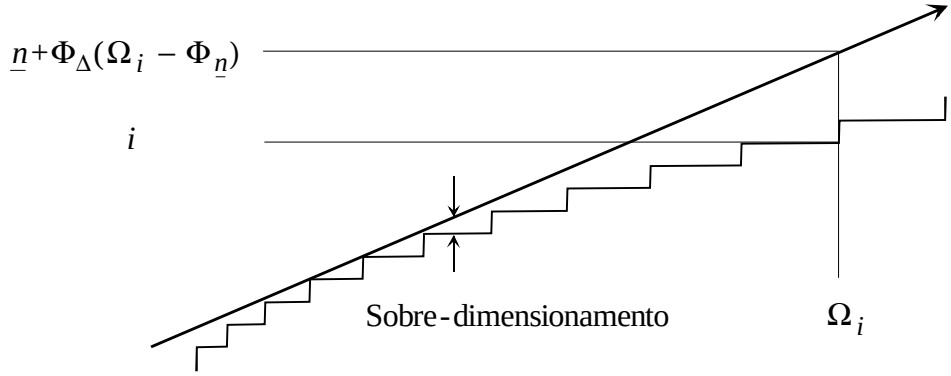


Figura 5-6 : Limites da diferença entre a base do degrau *i* de $ER^{-1}[\rho,P]$ e a semi-recta

O método escolhido para a determinação dos valores de $\Phi_{\underline{n}}$ e Φ_Δ baseia-se na minimização da soma das distâncias das bases de cada degrau da função original à semi-recta escolhida. Para um vértice genérico *i*, esta distância é ilustrada na Figura 5-6. Os valores de *i* a considerar são todos tais que $\underline{n} \le i \le \overline{n}$. Assim, a medida de qualidade adoptada para cada semi-recta possível é dada por:

$$Q = \sum_{i=\underline{n}}^{\overline{n}} \left(\underline{n} + \Phi_\Delta \left(\Omega_i - \Phi_{\underline{n}}\right) - i\right) \tag{17}$$

O método de cálculo de $\Omega_{\underline{n}}$ e Ω_Δ proposto é genericamente descrito pelo seguinte algoritmo:

Para *i* de $\underline{n}$ até $\overline{n}-1$ executar:

Calcular $\Phi_{\underline{n}}$ e Φ_Δ através de (16)

Calcular *Q* através de (17)

Guardar os valores de $\Phi_{\underline{n}}$ e Φ_Δ se o valor de *Q* for o melhor até agora

É possível que este cálculo possa ser efectuado por via analítica. Esta opção não foi abordada porque o tempo de cálculo da aproximação algorítmica é desprezável quando comparado com o tempo de cálculo da resolução dos modelos de programação resultantes (introduzidos na secção seguinte).

5.2 Modelos de Dimensionamento Uni-Horário

Considere-se uma rede com um conjunto de nós definido por N e um conjunto de arestas definido por A. Os elementos de A são representados por $\{i,j\}$. O elemento (i,j) representa a aresta $\{i,j\}$ no sentido de i para j.

O conjunto dos serviços ponto-multiponto suportados pela rede é definido pelo conjunto S. Para cada serviço $s \in S$ são definidos os conjuntos de nós S_u (nós de tráfego) e $\underline{S_u}$ (nós de trânsito). Os nós de tráfego são nós origem e destino de tráfego submetido à rede. Uma rede ATM suporta serviços ponto-multiponto através de ligações ponto-multiponto. O tráfego de um serviço s é definido por um conjunto de classes de tráfego F_s em que cada classe f é definida por:

- Um nó origem $k \in S_u$
- Um conjunto de nós destino $r \in S_u \setminus \{k\}$
- Uma intensidade de tráfego ρ_k^{sr} oferecida à rede do nó k para os nós r
- Um conjunto de variáveis binárias x_{ij}^{skr}, uma por cada arco (i,j) da rede, que representa a árvore de estabelecimento de ligação desta classe de tráfego na solução final.

Esta nomenclatura difere da nomenclatura apresentada na secção anterior em dois aspectos. Nos modelos de dimensionamento aqui propostos, os VPCs não são explicitamente modelizados. Assim, deixa de ser necessário o uso da letra k para os representar e as variáveis x das classes de tráfego passam a aplicar-se aos arcos e não aos VPCs como anteriormente. Para além disto, a letra o usada anteriormente para designar o nó origem de uma classe de tráfego confunde-se em algumas fórmulas com o valor zero. Assim, optou-se por usar nos modelos aqui propostos a letra k em vez da letra o para designar o nó origem de cada classe de tráfego.

O modelo de dimensionamento da rede tem de contemplar um conjunto de restrições que garanta que, para cada serviço s, a união das arestas para as quais todas as variáveis x_{ij}^{skr} são 1 seja uma árvore contendo pelo menos todos os nós pertencentes a S_u. Para o conseguir, começamos por introduzir as variáveis x_{ij}^{skd} que indicam o trajecto do nó origem $k \in S_u$ para o nó destino $d \in S_u$, incluído nas árvores de estabelecimento de ligação das classes com origem em k e cujo conjunto de destinos r contenha d. Usando as restrições seguintes:

$$\sum_{j \in N:\{i,j\} \in A} \left(x_{kj}^{skd} - x_{jk}^{skd} \right) = 1 \qquad , s \in S \, , k \in S_u \, , d \in S_u \setminus \{k\} \qquad (18a)$$

$$\sum_{j \in N:\{i,j\} \in A} \left(x_{jd}^{skd} - x_{dj}^{skd} \right) = 1 \qquad , s \in S \, , k \in S_u \, , d \in S_u \setminus \{k\} \qquad (18b)$$

$$\sum_{j \in N:\{i,j\} \in A} \left(x_{ij}^{skd} - x_{ji}^{skd} \right) = 0 \qquad , s \in S \, , k \in S_u \, , d \in S_u \setminus \{k\} \, , i \neq k \text{ e } i \neq d \qquad (18c)$$

garante-se que as variáveis x_{ij}^{skd} incluem um trajecto entre todos os pares de nós (k,d) pertencentes a S_u. Considere o conjunto das combinações r de nós pertencentes a $S_u \setminus \{k\}$ para as quais existe tráfego oferecido no nó $k \in S_u$ dado por R_k^s. Com as restrições:

$$x_{ij}^{skd} \leq x_{ij}^{skr} \qquad , s \in S \, , k \in S_u \, , r \in R_k^s : d \in r \, , (i,j) \in A \qquad (18d)$$

garante-se uma solução conexa para as variáveis x_{ij}^{skr}, ou seja, para todas as classes de tráfego, as respectivas variáveis x_{ij}^{skr} incluem um trajecto do seu nó origem para cada um dos seus nós destino.

As restrições apresentadas até agora ainda não garantem que as variáveis x_{ij}^{skr} definam uma árvore direccionada de k para os nós pertencentes a r. Para o conseguir, introduz-se um novo conjunto de variáveis:

p_{ij}^{sk} Variável binária: o valor 1 significa que a árvore de serviço do nó origem $k \in S_u$ para o serviço s passa no arco (i,j)

Conforme foi especificado na secção 5.1.1, a árvore de serviço do nó origem $k \in S_u$ para o serviço s é a árvore direccionada da origem para todos os outros nós de tráfego (pertencentes a $S_u \setminus \{k\}$) que define as árvores de estabelecimento de ligação da origem para qualquer das combinações de destinos $r \in R_k^s$. A introdução das restrições:

$$x_{ij}^{skd} \leq p_{ij}^{sk} \qquad , s \in S \, , k \in S_u \, , d \in S_u \setminus \{k\} \, , (i,j) \in A \qquad (18e)$$

em conjunto com as restrições (18d) garantem que as variáveis p_{ij}^{sk} incluem os arcos de das árvores de estabelecimento de ligação (definidas pelas variáveis x_{ij}^{skr}) de todas as classes de tráfego com origem em k. são a união de todos os trajectos de estabelecimento de ligação cuja origem é o nó k para o serviço s. Agora é apenas necessário garantir que as variáveis p_{ij}^{sk} definem uma árvore da origem k para todos os outros nós. Este objectivo é conseguido com a introdução das restrições seguintes:

$$\sum_{j \in N:\{i,j\} \in A} p_{ji}^{sk} \leq 1 \qquad , s \in S , k \in S_u , i \in N \qquad (18f)$$

As restrições (18f) impõem que o número de arcos de entrada em qualquer nó da rede é no máximo um pelo que os trajectos do mesmo nó origem *k* para diferentes destinos *d* (dados pelas variáveis x_{ij}^{skd}) não se podem cruzar entre si.

É necessário também garantir que nos nós de trânsito (nós pertencentes a $\underline{S_u}$) a árvore de serviço não tenha ramificações para que a solução final possa ser implementada com VPCs ponto-a-ponto apenas entre nós de tráfego. Para isso, é suficiente garantir que o número de arcos de saída destes nós é no máximo um através da introdução das restrições seguintes:

$$\sum_{j \in N:\{i,j\} \in A} p_{ij}^{sk} \leq 1 \qquad , s \in S , k \in S_u , i \in \underline{S_u} \qquad (18g)$$

Assim, as restrições introduzidas até agora garantem que para cada nó de serviço $k \in S_u$, a combinação das árvores de estabelecimento de ligação (dada pelas variáveis x_{ij}^{skr}) define umaa árvore de serviço (dada pelas variáveis p_{ij}^{sk}). É necessário ainda que, para cada serviço *s*, o conjunto de arestas usado por todos as árvores de serviço (dadas pelas variáveis p_{ij}^{sk} referentes a cada nó $k \in S_u$) formem uma árvore de suporte do serviço *s*. Este objectivo é conseguido introduzindo as variáveis seguintes:

$z_{\{ij\}}^{s}$ Variável binária de árvore de suporte: o valor 1 significa que a aresta $\{i,j\}$ pertence à árvore de suporte do serviço *s*

e as restrições seguintes:

$$p_{ij}^{sk} + p_{ji}^{sk} = z_{\{ij\}}^{s} \qquad , s \in S , k \in S_u , \{i,j\} \in A \qquad (18h)$$

Pelas restrições (18h), se a árvore de serviço para um nó origem *k* usa a aresta $\{i,j\}$, então a variável $z_{\{ij\}}^{s}$ é um e, por conseguinte, esta aresta também é usada pelas árvores de serviço de todos os nós $k \in S_u$ do mesmo serviço *s*.

Finalmente, é ainda necessário garantir, conforme foi abordado na secção 5.1.1, que o diâmetro da rede lógica seja limitado a um valor máximo de M_s para cada serviço *s*. Para isso é suficiente garantir que entre qualquer par de nós de tráfego (*k*,*d*) o número de arcos de saída dos nós pertencentes a S_u não é superior a M_s:

$$\sum_{i \in S_u} \sum_{j \in N : \{i,j\} \in A} x_{ji}^{skd} \leq M_s \qquad , s \in S \, , k \in S_u \, , d \in S_u \setminus \{k\} \qquad (18i)$$

Em conclusão, o conjunto de restrições (18a-i) garante que a solução para as variáveis x_{ij}^{skr}, referentes a todas as classes de tráfego de um serviço *s*, define uma rede lógica, dada pelas variáveis *z*, com uma topologia em árvore cujo diâmetro não ultrapassa o valor M_s. Se a estas restrições adicionarmos as restrições relativas aos modelos apresentados nas secção anterior, nomeadamente ao modelo binário:

$$\sum_{k \in S_u} \sum_{r \in R_k^s} \rho_k^{sr} \cdot x_{ij}^{skr} \leq \sum_{n=\underline{n}_s}^{\bar{n}_s} \Omega_n^s \cdot u_{ij}^{sn} \qquad , s \in S \, , (i,j) \in A \qquad (18j)$$

$$\sum_{n=\underline{n}_s}^{\bar{n}_s} u_{ij}^{sn} \leq 1 \qquad , s \in S \, , (i,j) \in A \qquad (18k)$$

ou ao modelo real:

$$\sum_{k \in S_u} \sum_{r \in R_k^s} \rho_k^{sr} \cdot x_{ij}^{skr} \leq \Phi_{\underline{n}}^s \cdot u_{ij}^s + \Phi_{\Delta}^s \cdot u_{ij}^{s\Delta} \, , s \in S \, , (i,j) \in A \qquad (18j')$$

$$u_{ij}^s \geq 0.0001 \cdot u_{ij}^{s\Delta} \qquad , s \in S \, , (i,j) \in A \qquad (18k')$$

obtemos um conjunto de restrições que simultaneamente garante uma solução para as variáveis x_{ij}^{skr} que segue uma topologia em árvore e uma solução para as variáveis *u* que cumpre com os requisitos de probabilidade de bloqueio máxima para os serviços suportados. As restrições (18j-h) e (18j'-k') são formalmente equivalentes às anteriormente apresentadas mas foi introduzido um índice adicional (o índice *s*) em todos os parâmetros e variáveis relativos ao serviço a que se referem.

Finalmente, falta apenas determinar um conjunto de restrições que garanta que a rede fisica tem capacidade para suportar as redes lógicas a configurar. Assim, considerem-se os parâmetros habituais que definem a rede física:

T	Conjunto dos tipos *t* de ligações físicas possíveis de implementar entre os nós de rede
α_t	Largura de banda de uma ligação física do tipo *t*
$Y_{\{ij\}}^t$	Número máximo de ligações físicas do tipo *t* entre os nós *i* e *j*
$C_{\{ij\}}^t$	Custo associado à utilização de uma ligação do tipo *t* entre os nós *i* e *j*
$y_{\{ij\}}^t$	Variável inteira que determina o número de ligações físicas do tipo *t* entre os nós *i* e *j*

Considere-se primeiro o modelo binário da fórmula de ErlangB inversa. Considerando uma largura de banda efectiva para o serviço *s* dada por w_s, introduzem-se as restrições seguintes:

$$\sum_{s\in S}\sum_{n=\underline{n}_s}^{\bar{n}_s}\left(w_s \cdot n \cdot u_{ij}^{sn}\right) \le \sum_{t\in T}\alpha_t y_{\{ij\}}^t \qquad , (i,j) \in A \tag{18l}$$

$$y_{\{ij\}}^t \le Y_{\{ij\}}^t \qquad , \{i,j\} \in A\ , t \in T \tag{18m}$$

As restrições (18l) impõem um limite inferior à soma das capacidades das ligações físicas a instalar em cada aresta da rede dada pela soma das capacidades das redes lógicas que por ela passam. As restrições (18m) impõem que o número de ligações na solução final não é maior do que o seu valor máximo permitido.

Considere-se agora o modelo real da fórmula de ErlangB inversa. Neste caso, as restrições adequadas são:

$$\sum_{s\in S} w_s\left(\underline{n}_s \cdot u_{ij}^s + u_{ij}^{s\Delta}\right) \le \sum_{t\in T}\alpha_t y_{\{ij\}}^t \qquad , (i,j) \in A \tag{18l'}$$

$$y_{\{ij\}}^t \le Y_{\{ij\}}^t \qquad , \{i,j\} \in A\ , t \in T \tag{18m}$$

Para se perceber o lado esquerdo da restrição (18l'), relembre-se que a variável binária u_{ij}^s representa, se estiver a um, que um VPC da rede lógica de suporte ao serviço *s* que passa em (*i*,*j*) tem pelo menos uma capacidade de $\underline{n}_s$ ligações e a variável real $u_{ij}^{s\Delta}$ determina a capacidade adicional do VPC. Assim, a capacidade total, em número de ligações lógicas, da rede lógica que passa em (*i*,*j*) é dada por $\underline{n}_s \cdot u_{ij}^s + u_{ij}^{s\Delta}$.

As sub-secções seguintes apresentam os dois modelos completos considerados para o dimensionamento de redes ATM uni-horário de suporte a redes lógicas em árvore. O modelo A baseia-se no modelo binário para a aproximação da fórmula de ErlangB inversa e o modelo B baseia-se no modelo real para esta aproximação.

5.2.1 Modelo A de Dimensionamento Uni-Horário

O modelo de Programação Inteira que se apresenta a seguir formula um problema de optimização cuja resolução escolhe a combinação de árvores de suporte (uma para cada serviço) que resultam na rede física de menor custo.

N Conjunto de nós da rede

A Conjunto de arestas $\{i,j\}$ da rede: (i,j) representa a aresta $\{i,j\}$ no sentido de i para j e (j,i) representa o sentido inverso

T Conjunto dos tipos t de ligações físicas possíveis de implementar entre os nós de rede

α_t Largura de banda de uma ligação física do tipo t

$Y^t_{\{ij\}}$ Número máximo de ligações físicas do tipo t entre os nós i e j

$C^t_{\{ij\}}$ Custo associado à utilização de uma ligação do tipo t entre os nós i e j

S Conjunto de serviços ponto-multiponto s suportados pela rede

M_s Diâmetro máximo da rede lógica de suporte ao serviço s

S_u Conjunto dos nós de tráfego do serviço s

$\underline{S}_u$ Conjunto dos nós de trânsito do serviço s ($S_u \cup \underline{S}_u = N$ e $S_u \cap \underline{S}_u = \varnothing$)

Ω^s_n Máximo tráfego (em Erlangs) suportado por um VPC com capacidade para n ligações de suporte ao serviço s

$\underline{n}_s$ Número mínimo de ligações de qualquer VPC de suporte ao serviço s

$\overline{n}_s$ Número máximo de ligações de qualquer VPC de suporte ao serviço s

R^s_k Conjunto das combinações r de nós pertencentes a $S_u \setminus \{k\}$ para as quais existe tráfego oferecido no nó $k \in S_u$

ρ^{sr}_k Tráfego oferecido (em Erlangs) pelo serviço s no nó origem $k \in S_u$ para a combinação de nós destino $r \in R^s_k$

w_s Largura de banda de uma ligação do serviço s

$y^t_{\{ij\}}$ Variável inteira que determina o número de ligações físicas do tipo t entre os nós i e j

p^{sk}_{ij} Variável binária: o valor 1 significa que, para o serviço s, a árvore de serviço com origem no nó $k \in S_u$ passa no arco (i,j)

x^{skr}_{ij} Variável binária: o valor 1 significa que, para o serviço s, a árvore de estabelecimento de ligação do nó $k \in S_u$ para a combinação de nós destino $r \in R^s_k$ passa no arco (i,j)

x^{skd}_{ij} Variável binária: o valor 1 significa que, para o serviço s, o trajecto do nó $k \in S_u$ para o nó $d \in S_u$ incluído nas árvores de estabelecimento de ligação das classes com origem em k e cujo conjunto de destinos r contenha d, passa no arco (i,j)

$z^s_{\{ij\}}$ Variável binária de árvore de suporte: o valor 1 significa que a aresta $\{i,j\}$ pertence à árvore de suporte do serviço s

u_{ij}^{sn} Variável binária de capacidade: o valor 1 significa que o arco (*i*,*j*) da árvore de suporte do serviço *s* tem capacidade para *n* ligações simultâneas

Minimizar $\sum_{\{i,j\}\in A}\sum_{t\in T} C_{\{ij\}}^{t} y_{\{ij\}}^{t}$

Sujeito a:

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{kj}^{skd} - x_{jk}^{skd}\right)=1 \qquad , s\in S\,,\, k\in S_u\,,\, d\in S_u\setminus\{k\} \qquad (18a)$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{jd}^{skd} - x_{dj}^{skd}\right)=1 \qquad , s\in S\,,\, k\in S_u\,,\, d\in S_u\setminus\{k\} \qquad (18b)$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{ij}^{skd} - x_{ji}^{skd}\right)=0 \qquad , s\in S\,,\, k\in S_u\,,\, d\in S_u\setminus\{k\}\,,\, i\neq k \text{ e } i\neq d \qquad (18c)$$

$$x_{ij}^{skd} \le x_{ij}^{skr} \qquad , s\in S\,,\, k\in S_u\,,\, r\in R_k^s : d\in r\,,(i,j)\in A \qquad (18d)$$

$$x_{ij}^{skd} \le p_{ij}^{sk} \qquad , s\in S\,,\, k\in S_u\,,\, d\in S_u\setminus\{k\}\,,\, (i,j)\in A \qquad (18e)$$

$$\sum_{j\in N:\{i,j\}\in A} p_{ji}^{sk} \le 1 \qquad , s\in S\,,\, k\in S_u\,,\, i\in N \qquad (18f)$$

$$\sum_{j\in N:\{i,j\}\in A} p_{ij}^{sk} \le 1 \qquad , s\in S\,,\, k\in S_u\,,\, i\in \underline{S}_u \qquad (18g)$$

$$p_{ij}^{sk} + p_{ji}^{sk} = z_{\{ij\}}^{s} \qquad , s\in S\,,\, k\in S_u\,,\, \{i,j\}\in A \qquad (18h)$$

$$\sum_{i\in S_u}\sum_{j\in N:\{i,j\}\in A} x_{ji}^{skd} \le M_s \qquad , s\in S\,,\, k\in S_u\,,\, d\in S_u\setminus\{k\} \qquad (18i)$$

$$\sum_{k\in S_u}\sum_{r\in R_k^s} \rho_k^{sr}\cdot x_{ij}^{skr} \le \sum_{n=\underline{n}_s}^{\bar{n}_s} \Omega_n^s \cdot u_{ij}^{sn} \qquad , s\in S\,,\, (i,j)\in A \qquad (18j)$$

$$\sum_{n=\underline{n}_s}^{\bar{n}_s} u_{ij}^{sn} \le 1 \qquad , s\in S\,,\, (i,j)\in A \qquad (18k)$$

$$\sum_{s\in S}\sum_{n=\underline{n}_s}^{\bar{n}_s}\left(w_s\cdot n\cdot u_{ij}^{sn}\right) \le \sum_{t\in T}\alpha_t y_{\{ij\}}^{t} \qquad , (i,j)\in A \qquad (18l)$$

$$y_{\{ij\}}^{t} \le Y_{\{ij\}}^{t} \qquad , \{i,j\}\in A\,,\, t\in T \qquad (18m)$$

$x_{ij}^{skr} \in \{0,1\}$, $p_{ij}^{sk} \in \{0,1\}$, $z_{\{ij\}}^{s} \in \{0,1\}$, $u_{ij}^{sn} \in \{0,1\}$, $y_{\{ij\}}^{t}$ inteiro não negativo

De notar que o trajecto e a capacidade dos VPCs que constituem cada árvore de suporte não são directamente modelizados na formulação apresentada. No entanto, estas grandezas estão indirectamente representadas pelas variáveis $z^s_{\{ij\}}$ e u^{sn}_{ij}. Cada conjunto de arestas da árvore de suporte $\{\{i,j\}: z^s_{\{ij\}} = 1\}$ que definam um trajecto entre dois nós de tráfego do serviço *s*, define o trajecto de um VPC. Para cada um dos VPCs, a grandeza:

$$\sum_{n=\underline{n}_s}^{\bar{n}_s} \left(w_s \cdot n \cdot u^{sn}_{ij} \right)$$

define a capacidade do VPC no arco (*i*,*j*) por onde passa.

5.2.2 Modelo B de Dimensionamento Uni-Horário

O modelo de Programação Inteira Mista que se apresenta a seguir formula um problema de optimização cuja resolução escolhe a combinação de árvores de suporte (uma para cada serviço) que resultam na rede física de menor custo.

N Conjunto de nós da rede

A Conjunto de arestas $\{i,j\}$ da rede: (i,j) representa a aresta $\{i,j\}$ no sentido de *i* para *j* e (j,i) representa o sentido inverso

T Conjunto dos tipos *t* de ligações físicas possíveis de implementar entre os nós de rede

α_t Largura de banda de uma ligação física do tipo *t*

$Y^t_{\{ij\}}$ Número máximo de ligações físicas do tipo *t* entre os nós *i* e *j*

$C^t_{\{ij\}}$ Custo associado à utilização de uma ligação do tipo *t* entre os nós *i* e *j*

S Conjunto de serviços ponto-multiponto *s* suportados pela rede

M_s Diâmetro máximo da rede lógica de suporte ao serviço *s*

S_u Conjunto dos nós de tráfego do serviço *s*

$\underline{S}_u$ Conjunto dos nós de trânsito do serviço *s* ($S_u \cup \underline{S}_u = N$ e $S_u \cap \underline{S}_u = \varnothing$)

$\underline{n}_s$ Número mínimo de ligações de qualquer VPC de suporte ao serviço *s*

$\Phi^s_{\underline{n}}$ Máximo tráfego (em Erlangs) suportado por um VPC de suporte ao serviço *s* com capacidade para $\underline{n}_s$ ligações simultâneas

Φ^s_{Δ} Capacidade de tráfego (em Erlangs) adicional suportado por um VPC de suporte ao serviço *s* para cada ligação acima do valor mínimo de $\underline{n}_s$ ligações

R^s_k Conjunto das combinações *r* de nós pertencentes a $S_u \setminus \{k\}$ para as quais existe tráfego oferecido no nó $k \in S_u$

ρ_k^{sr} Tráfego oferecido (em Erlangs) pelo serviço s no nó origem $k \in S_u$ para a combinação de nós destino $r \in R_k^s$

w_s Largura de banda de uma ligação do serviço s

$y_{\{ij\}}^t$ Variável inteira que determina o número de ligações físicas do tipo t entre os nós i e j

p_{ij}^{sk} Variável binária: o valor 1 significa que, para o serviço s, a árvore de serviço com origem no nó $k \in S_u$ passa no arco (i,j)

x_{ij}^{skr} Variável binária: o valor 1 significa que, para o serviço s, a árvore de estabelecimento de ligação do nó $k \in S_u$ para a combinação de nós destino $r \in R_k^s$ passa no arco (i,j)

x_{ij}^{skd} Variável binária: o valor 1 significa que, para o serviço s, o trajecto do nó $k \in S_u$ para o nó $d \in S_u$ incluído nas árvores de estabelecimento de ligação das classes com origem em k e cujo conjunto de destinos r contenha d, passa no arco (i,j)

$z_{\{ij\}}^s$ Variável binária de árvore de suporte: o valor 1 significa que a aresta $\{i,j\}$ pertence à árvore de suporte do serviço s

u_{ij}^s Variável binária de capacidade: o valor 1 significa que o arco (i,j) da árvore de suporte do serviço s tem capacidade para pelo menos $\underline{n}_s$ ligações simultâneas

$u_{ij}^{s\Delta}$ Variável real de capacidade cujo valor representa a capacidade adicional de ligações no arco (i,j) da árvore de suporte do serviço s acima das $\underline{n}_s$ ligações simultâneas

Minimizar $$\sum_{\{i,j\}\in A}\sum_{t\in T} C_{\{ij\}}^t y_{\{ij\}}^t$$

Sujeito a:

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{kj}^{skd} - x_{jk}^{skd}\right)=1 \qquad , s\in S\,,\, k\in S_u\,,\, d\in S_u\setminus\{k\} \qquad (18a)$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{jd}^{skd} - x_{dj}^{skd}\right)=1 \qquad , s\in S\,,\, k\in S_u\,,\, d\in S_u\setminus\{k\} \qquad (18b)$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{ij}^{skd} - x_{ji}^{skd}\right)=0 \qquad , s\in S\,,\, k\in S_u\,,\, d\in S_u\setminus\{k\}\,,\, i\neq k \text{ e } i\neq d \qquad (18c)$$

$$x_{ij}^{skd} \le x_{ij}^{skr} \qquad , s\in S\,,\, k\in S_u\,,\, r\in R_k^s : d\in r\,,(i,j)\in A \qquad (18d)$$

$$x_{ij}^{skd} \le p_{ij}^{sk} \qquad , s\in S\,,\, k\in S_u\,,\, d\in S_u\setminus\{k\}\,,\, (i,j)\in A \qquad (18e)$$

$$\sum_{j\in N:\{i,j\}\in A} p_{ji}^{sk} \le 1 \qquad , s\in S\,,\, k\in S_u\,,\, i\in N \qquad (18f)$$

$$\sum_{j \in N:\{i,j\} \in A} p_{ij}^{sk} \leq 1 \qquad , s \in S\ , k \in S_u\ , i \in \underline{S}_u \qquad (18g)$$

$$p_{ij}^{sk} + p_{ji}^{sk} = z_{\{ij\}}^{s} \qquad , s \in S\ , k \in S_u\ , \{i,j\} \in A \qquad (18h)$$

$$\sum_{i \in S_u} \sum_{j \in N:\{i,j\} \in A} x_{ji}^{skd} \leq M_s \qquad , s \in S\ , k \in S_u\ , d \in S_u \setminus \{k\} \qquad (18i)$$

$$\sum_{k \in S_u} \sum_{r \in R_k^s} \rho_k^{sr} \cdot x_{ij}^{skr} \leq \Phi_{\underline{n}}^{s} \cdot u_{ij}^{s} + \Phi_{\Delta}^{s} \cdot u_{ij}^{s\Delta} \qquad , s \in S\ , (i,j) \in A \qquad (18j')$$

$$u_{ij}^{s} \geq 0.0001 \cdot u_{ij}^{s\Delta} \qquad , s \in S\ , (i,j) \in A \qquad (18k')$$

$$\sum_{s \in S} w_s \left(\underline{n}_s \cdot u_{ij}^{s} + u_{ij}^{s\Delta} \right) \leq \sum_{t \in T} \alpha_t y_{\{ij\}}^{t} \qquad , (i,j) \in A \qquad (18l')$$

$$y_{\{ij\}}^{t} \leq Y_{\{ij\}}^{t} \qquad , \{i,j\} \in A\ , t \in T \qquad (18m)$$

$x_{ij}^{skr} \in \{0,1\}$, $p_{ij}^{sk} \in \{0,1\}$, $z_{\{ij\}}^{s} \in \{0,1\}$, $u_{ij}^{s} \in \{0,1\}$

$u_{ij}^{s\Delta}$ real não negativo $y_{\{ij\}}^{t}$ inteiro não negativo

De notar que o trajecto e a capacidade dos VPCs que constituem cada árvore de suporte não são directamente modelizados na formulação apresentada. No entanto, estas grandezas estão indirectamente representadas pelas variáveis $z_{\{ij\}}^{s}$, u_{ij}^{s} e $u_{ij}^{s\Delta}$. Cada conjunto de arestas da árvore de suporte $\{\{i,j\}: z_{\{ij\}}^{s}=1\}$ que definam um trajecto entre dois nós de tráfego do serviço *s*, define o trajecto de um VPC. Para cada um dos VPCs, a grandeza:

$$w_s \left(\underline{n}_s \cdot u_{ij}^{s} + u_{ij}^{s\Delta} \right)$$

define a capacidade do VPC no arco (*i*,*j*) por onde passa.

5.3 Reformulações dos Modelos de Dimensionamento Uni-Horário

A eficiência da resolução de um modelo de Programação Inteira ou Programação Inteira Mista através do algoritmo de *branch-and-bound* é dependente da sua complexidade (número de variáveis e número de restrições) e do valor da solução da sua relaxação linear (ver Anexo C). Considera-se que os modelos apresentados nas secções anteriores são os modelos naturais do problema de dimensionamento abordado. Quando é

possível reformular um modelo (quer por introdução ou alteração de variáveis, quer por introdução ou alteração de restrições) de modo a que o espaço de soluções possíveis da relaxação linear seja diminuído sem que nenhuma das soluções inteiras do problema original seja eliminada, então o valor da relaxação linear do problema reformulado é necessariamente maior ou igual (para um problema de minimização) do que o do modelo original. A eficiência do modelo reformulado é tanto melhor quanto maior for o valor da relaxação linear e é tanto pior quanto maior for a complexidade introduzida. Nos modelos apresentados na secção anterior, foram identificadas reformulações válidas que a seguir se descrevem e cuja eficiência será abordada na secção dedicada a resultados computacionais.

Ao contrário das secções anteriores, considere-se primeiro o modelo B baseado no modelo real da função ErlangB inversa. Sempre que para um determinado serviço *s* e num determinado arco (*i*,*j*) uma das variáveis x_{ij}^{skr} seja um, então pelas restrições (18j') e (18k') a variável u_{ij}^{s} também é um. Por outro lado, a variável x_{ij}^{skr} = 1 impõe, pelas restrições (18a-h), que a aresta {*i*,*j*} pertença à árvore de suporte de *s* (como foi discutido na secção anterior) e, portanto, a variável $z_{\{ij\}}^{s}$ também é um. Como nos cenários de tráfego de interesse (secção 2.4.2), existe tráfego nos dois sentidos de cada aresta pertencente à árvore de suporte da cada serviço *s*, então o valor da variável $z_{\{ij\}}^{s}$ é igual ao valor das variáveis u_{ij}^{s} em ambos os sentidos da aresta {*i*,*j*}. Assim, uma reformulação possível para o modelo B é a eliminação da variável u_{ij}^{s} e a sua substituição pela variável $z_{\{ij\}}^{s}$. Esta reformulação implementa-se substituindo as restrições (18j'), (18k') e (18l') pelas restrições:

$$\sum_{k\in S_u}\sum_{r\in R_k^s}\rho_k^{sr}\cdot x_{ij}^{skr} \le \Phi_{\underline{n}}^{s}\cdot z_{\{ij\}}^{s} + \Phi_{\Delta}^{s}\cdot u_{ij}^{s\Delta} \qquad , s\in S\,,\,(i,j)\in A \tag{19j'}$$

$$\sum_{s\in S} w_s\left(\underline{n}_s\cdot z_{\{ij\}}^{s} + u_{ij}^{s\Delta}\right) \le \sum_{t\in T}\alpha_t y_{\{ij\}}^{t} \qquad , (i,j)\in A \tag{19l'}$$

De notar que as restrições (18k') foram eliminadas pois a sua função era garantir que, em cada serviço *s* e cada arco (*i*,*j*), a variável u_{ij}^{s} seria um para que a variável real $u_{ij}^{s\Delta}$ pudesse ser maior que zero. Com a eliminação das variáveis u_{ij}^{s}, também estas restrições podem ser eliminadas.

Considere-se agora o modelo A proposto na secção anterior. A reformulação que se propõe baseia-se no mesmo raciocínio da reformulação proposta para o modelo B mas a

sua implementação é mais complexa. O objectivo é substituir a variável binária u_{ij}^{sn} para $n = \underline{n}_s$ pela variável $z_{\{ij\}}^{s}$. No entanto, para que seja possível fazê-lo na restrição (18j), é preciso reformular os valores de Ω_n^s. Seja $\Omega_{\underline{n}}^s$ o valor de Ω_n^s para $n = \underline{n}_s$. Considere que para o serviço *s*, $\underline{\Omega}_n^s$ representa a diferença entre o tráfego suportado por *n* ligações e o tráfego suportado por $\underline{n}_s$ ligações, ou seja:

$$\underline{\Omega}_n^s = \Omega_n^s - \Omega_{\underline{n}}^s \qquad , s \in S \, , n > \underline{n}_s$$

Com a introdução das grandezas $\underline{\Omega}_n^s$, é possível substituir as restrições (18j), (18k) e (18l) pelas restrições:

$$\sum_{k \in S_u} \sum_{r \in R_k^s} \rho_k^{sr} \cdot x_{ij}^{skr} \le \Omega_{\underline{n}}^s \cdot z_{\{ij\}}^s + \sum_{n=\underline{n}_s+1}^{\bar{n}_s} \underline{\Omega}_n^s \cdot u_{ij}^{sn} \qquad , s \in S \, , (i,j) \in A \qquad (19j)$$

$$\sum_{n=\underline{n}_s+1}^{\bar{n}_s} u_{ij}^{sn} \le z_{\{ij\}}^s \qquad , s \in S \, , (i,j) \in A \qquad (19k)$$

$$\sum_{s \in S} w_s \left(\underline{n}_s \cdot z_{\{ij\}}^s + \sum_{n=\underline{n}_s+1}^{\bar{n}_s} \left[\left(n - \underline{n}_s \right) \cdot u_{ij}^{sn} \right] \right) \le \sum_{t \in T} \alpha_t y_{\{ij\}}^t \qquad , (i,j) \in A \qquad (19l)$$

em que a variável u_{ij}^{sn} , para $n = \underline{n}_s$ é substituída pela variável $z_{\{ij\}}^s$ nas restrições (19j) e (19l). A substituição da constante 1 pela variável $z_{\{ij\}}^s$ nas restrições (19k) funciona como um corte no problema original porque elimina algumas soluções possíveis da relaxação linear sem que se tenha eliminado nenhuma das soluções inteiras.

Em conclusão, as reformulações propostas diminuem ligeiramente a complexidade dos modelos originais. Em ambos os modelos, as reformulações representam a diminuição do número total de variáveis e no modelo B, representa. também a diminuição do número total de restrições. As sub-secções seguintes apresentam os dois modelos reformulados completos considerados para o dimensionamento de redes ATM uni-horário de suporte a redes lógicas em árvore. A eficiência das reformulações propostas será analisada na secção dedicada a resultados computacionais.

5.3.1 Modelo A Reformulado de Dimensionamento Uni-Horário

N Conjunto de nós da rede

A Conjunto de arestas {*i*,*j*} da rede: (*i*,*j*) representa a aresta {*i*,*j*} no sentido de *i* para *j* e (*j*,*i*) representa o sentido inverso

T Conjunto dos tipos *t* de ligações físicas possíveis de implementar entre os nós de rede

α_t Largura de banda de uma ligação física do tipo *t*

$Y_{\{ij\}}^t$ Número máximo de ligações físicas do tipo *t* os nós *i* e *j*

$C_{\{ij\}}^t$ Custo associado à utilização de uma ligação do tipo *t* entre os nós *i* e *j*

S Conjunto de serviços ponto-multiponto *s* suportados pela rede

M_s Diâmetro máximo da rede lógica de suporte ao serviço *s*

S_u Conjunto dos nós de tráfego do serviço *s*

$\underline{S}_u$ Conjunto dos nós de trânsito do serviço *s* ($S_u \cup \underline{S}_u = N$ e $S_u \cap \underline{S}_u = \varnothing$)

$\underline{n}_s$ Número mínimo de ligações de qualquer VPC de suporte ao serviço *s*

$\overline{n}_s$ Número máximo de ligações de qualquer VPC de suporte ao serviço *s*

$\Omega_{\underline{n}}^s$ Máximo tráfego (em Erlangs) suportado por um VPC de suporte ao serviço *s* com capacidade para $\underline{n}_s$ ligações

$\underline{\Omega}_n^s$ Capacidade de tráfego (em Erlangs) adicional suportado por um VPC de suporte ao serviço *s* com capacidade para *n* ligações. Este valor é a diferença entre o tráfego suportado por *n* ligações e o tráfego suportado por $\underline{n}_s$ ligações

R_k^s Conjunto das combinações *r* de nós pertencentes a $S_u \setminus \{k\}$ para as quais existe tráfego oferecido no nó $k \in S_u$

ρ_k^{sr} Tráfego oferecido (em Erlangs) pelo serviço *s* no nó origem $k \in S_u$ para a combinação de nós destino $r \in R_k^s$

w_s Largura de banda de uma ligação do serviço *s*

$y_{\{ij\}}^t$ Variável inteira que determina o número de ligações físicas do tipo *t* entre os nós *i* e *j*

p_{ij}^{sk} Variável binária: o valor 1 significa que, para o serviço *s*, a árvore de serviço com origem no nó $k \in S_u$ passa no arco (*i*,*j*)

x_{ij}^{skr} Variável binária: o valor 1 significa que, para o serviço *s*, a árvore de estabelecimento de ligação do nó $k \in S_u$ para a combinação de nós destino $r \in R_k^s$ passa no arco (*i*,*j*)

x_{ij}^{skd} Variável binária: o valor 1 significa que, para o serviço *s*, o trajecto do nó $k \in S_u$ para o nó $d \in S_u$ incluído nas árvores de estabelecimento de ligação das classes com origem em *k* e cujo conjunto de destinos *r* contenha *d*, passa no arco (*i*,*j*)

$z_{\{ij\}}^s$ Variável binária de árvore de suporte: o valor 1 significa que a aresta {*i*,*j*} pertence à árvore de suporte do serviço *s*

u_{ij}^{sn} Variável binária de capacidade: o valor 1 significa que o arco (*i*,*j*) da árvore de suporte do serviço *s* tem capacidade para *n* ligações simultâneas

Minimizar $\sum_{\{i,j\}\in A}\sum_{t\in T} C_{\{ij\}}^{t} y_{\{ij\}}^{t}$

Sujeito a:

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{kj}^{skd} - x_{jk}^{skd}\right)=1 \qquad , s\in S\,,\, k\in S_u\,,\, d\in S_u\setminus\{k\} \qquad (19a)$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{jd}^{skd} - x_{dj}^{skd}\right)=1 \qquad , s\in S\,,\, k\in S_u\,,\, d\in S_u\setminus\{k\} \qquad (19b)$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{ij}^{skd} - x_{ji}^{skd}\right)=0 \qquad , s\in S\,,\, k\in S_u\,,\, d\in S_u\setminus\{k\}\,,\, i\neq k \text{ e } i\neq d \qquad (19c)$$

$$x_{ij}^{skd} \le x_{ij}^{skr} \qquad , s\in S\,,\, k\in S_u\,,\, r\in R_k^s : d\in r\,,(i,j)\in A \qquad (19d)$$

$$x_{ij}^{skd} \le p_{ij}^{sk} \qquad , s\in S\,,\, k\in S_u\,,\, d\in S_u\setminus\{k\}\,,\, (i,j)\in A \qquad (19e)$$

$$\sum_{j\in N:\{i,j\}\in A} p_{ji}^{sk} \le 1 \qquad , s\in S\,,\, k\in S_u\,,\, i\in N \qquad (19f)$$

$$\sum_{j\in N:\{i,j\}\in A} p_{ij}^{sk} \le 1 \qquad , s\in S\,,\, k\in S_u\,,\, i\in \underline{S}_u \qquad (19g)$$

$$p_{ij}^{sk} + p_{ji}^{sk} = z_{\{ij\}}^{s} \qquad , s\in S\,,\, k\in S_u\,,\, \{i,j\}\in A \qquad (19h)$$

$$\sum_{i\in S_u}\sum_{j\in N:\{i,j\}\in A} x_{ji}^{skd} \le M_s \qquad , s\in S\,,\, k\in S_u\,,\, d\in S_u\setminus\{k\} \qquad (19i)$$

$$\sum_{k\in S_u}\sum_{r\in R_k^s}\rho_k^{sr}\cdot x_{ij}^{skr} \le \Omega_{\underline{n}}^{s}\cdot z_{\{ij\}}^{s} + \sum_{n=\underline{n}_s+1}^{\bar{n}_s}\underline{\Omega}_n^{s}\cdot u_{ij}^{sn} \qquad , s\in S\,,\, (i,j)\in A \qquad (19j)$$

$$\sum_{n=\underline{n}_s+1}^{\bar{n}_s} u_{ij}^{sn} \le z_{\{ij\}}^{s} \qquad , s\in S\,,\, (i,j)\in A \qquad (19k)$$

$$\sum_{s\in S} w_s\left(\underline{n}_s\cdot z_{\{ij\}}^{s} + \sum_{n=\underline{n}_s+1}^{\bar{n}_s}\left[(n-\underline{n}_s)\cdot u_{ij}^{sn}\right]\right) \le \sum_{t\in T}\alpha_t y_{\{ij\}}^{t} \qquad , (i,j)\in A \qquad (19l)$$

$$y_{\{ij\}}^{t} \le Y_{\{ij\}}^{t} \qquad , \{i,j\}\in A\,,\, t\in T \qquad (19m)$$

$x_{ij}^{skr} \in \{0,1\}$, $p_{ij}^{sk} \in \{0,1\}$, $z_{\{ij\}}^{s} \in \{0,1\}$, $u_{ij}^{sn} \in \{0,1\}$, $y_{\{ij\}}^{t}$ inteiro não negativo

5.3.2 Modelo B Reformulado de Dimensionamento Uni-Horário

N Conjunto de nós da rede

A Conjunto de arestas $\{i,j\}$ da rede: (i,j) representa a aresta $\{i,j\}$ no sentido de i para j e (j,i) representa o sentido inverso

T Conjunto dos tipos t de ligações físicas possíveis de implementar entre os nós de rede

α_t Largura de banda de uma ligação física do tipo t

$Y_{\{ij\}}^t$ Número máximo de ligações físicas do tipo t entre os nós i e j

$C_{\{ij\}}^t$ Custo associado à utilização de uma ligação do tipo t entre os nós i e j

S Conjunto de serviços ponto-multiponto s suportados pela rede

M_s Diâmetro máximo da rede lógica de suporte ao serviço s

S_u Conjunto dos nós de tráfego do serviço s

$\underline{S}_u$ Conjunto dos nós de trânsito do serviço s ($S_u \cup \underline{S}_u = N$ e $S_u \cap \underline{S}_u = \varnothing$)

$\underline{n}_s$ Número mínimo de ligações de qualquer VPC de suporte ao serviço s

$\Phi_{\underline{n}}^s$ Máximo tráfego (em Erlangs) suportado por um VPC de suporte ao serviço s com capacidade para $\underline{n}_s$ ligações simultâneas

Φ_{Δ}^s Capacidade de tráfego (em Erlangs) adicional suportado por um VPC de suporte ao serviço s para cada ligação acima do valor mínimo de $\underline{n}_s$ ligações

R_k^s Conjunto das combinações r de nós pertencentes a $S_u \setminus \{k\}$ para as quais existe tráfego oferecido no nó $k \in S_u$

ρ_k^{sr} Tráfego oferecido (em Erlangs) pelo serviço s no nó origem $k \in S_u$ para a combinação de nós destino $r \in R_k^s$

w_s Largura de banda de uma ligação do serviço s

$y_{\{ij\}}^t$ Variável inteira que determina o número de ligações físicas do tipo t entre os nós i e j

p_{ij}^{sk} Variável binária: o valor 1 significa que, para o serviço s, a árvore de serviço com origem no nó $k \in S_u$ passa no arco (i,j)

x_{ij}^{skr} Variável binária: o valor 1 significa que, para o serviço s, a árvore de estabelecimento de ligação do nó $k \in S_u$ para a combinação de nós destino $r \in R_k^s$ passa no arco (i,j)

x_{ij}^{skd} Variável binária: o valor 1 significa que, para o serviço s, o trajecto do nó $k \in S_u$ para o nó $d \in S_u$ incluído nas árvores de estabelecimento de ligação das classes com origem em k e cujo conjunto de destinos r contenha d, passa no arco (i,j)

$z^s_{\{ij\}}$ Variável binária de árvore de suporte: o valor 1 significa que a aresta $\{i,j\}$ pertence à árvore de suporte do serviço s

$u^{s\Delta}_{ij}$ Variável real de capacidade cujo valor representa a capacidade adicional de ligações no arco (i,j) da árvore de suporte do serviço s acima das $\underline{n}_s$ ligações simultâneas

Minimizar $\sum_{\{i,j\}\in A}\sum_{t\in T} C^t_{\{ij\}} y^t_{\{ij\}}$

Sujeito a:

$$\sum_{j\in N:\{i,j\}\in A}\left(x^{skd}_{kj} - x^{skd}_{jk}\right)=1 \qquad , s\in S\,,\, k\in S_u\,,\, d\in S_u\setminus\{k\} \qquad (19a)$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x^{skd}_{jd} - x^{skd}_{dj}\right)=1 \qquad , s\in S\,,\, k\in S_u\,,\, d\in S_u\setminus\{k\} \qquad (19b)$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x^{skd}_{ij} - x^{skd}_{ji}\right)=0 \qquad , s\in S\,,\, k\in S_u\,,\, d\in S_u\setminus\{k\}\,,\, i\neq k \text{ e } i\neq d \qquad (19c)$$

$$x^{skd}_{ij} \le x^{skr}_{ij} \qquad , s\in S\,,\, k\in S_u\,,\, r\in R^s_k : d\in r\,,(i,j)\in A \qquad (19d)$$

$$x^{skd}_{ij} \le p^{sk}_{ij} \qquad , s\in S\,,\, k\in S_u\,,\, d\in S_u\setminus\{k\}\,,\, (i,j)\in A \qquad (19e)$$

$$\sum_{j\in N:\{i,j\}\in A} p^{sk}_{ji} \le 1 \qquad , s\in S\,,\, k\in S_u\,,\, i\in N \qquad (19f)$$

$$\sum_{j\in N:\{i,j\}\in A} p^{sk}_{ij} \le 1 \qquad , s\in S\,,\, k\in S_u\,,\, i\in \underline{S}_u \qquad (19g)$$

$$p^{sk}_{ij} + p^{sk}_{ji} = z^s_{\{ij\}} \qquad , s\in S\,,\, k\in S_u\,,\, \{i,j\}\in A \qquad (19h)$$

$$\sum_{i\in S_u}\sum_{j\in N:\{i,j\}\in A} x^{skd}_{ji} \le M_s \qquad , s\in S\,,\, k\in S_u\,,\, d\in S_u\setminus\{k\} \qquad (19i)$$

$$\sum_{k\in S_u}\sum_{r\in R^s_k} \rho^{sr}_k \cdot x^{skr}_{ij} \le \Phi^s_{\underline{n}} \cdot z^s_{\{ij\}} + \Phi^s_{\Delta}\cdot u^{s\Delta}_{ij} \qquad , s\in S\,,\, (i,j)\in A \qquad (19j')$$

$$\sum_{s\in S} w_s\left(\underline{n}_s \cdot z^s_{\{ij\}} + u^{s\Delta}_{ij}\right) \le \sum_{t\in T}\alpha_t y^t_{\{ij\}} \qquad , (i,j)\in A \qquad (19l')$$

$$y^t_{\{ij\}} \le Y^t_{\{ij\}} \qquad , \{i,j\}\in A\,,\, t\in T \qquad (19m)$$

$x^{skr}_{ij} \in \{0,1\}$, $p^{sk}_{ij} \in \{0,1\}$, $z^s_{\{ij\}} \in \{0,1\}$

$u^{s\Delta}_{ij}$ real não negativo $\qquad$ $y^t_{\{ij\}}$ inteiro não negativo

5.4 Modelos de Dimensionamento Multi-Horário

Os modelos de Programação Inteira e Programação Inteira Mista que se apresentam nas sub-secções seguintes são reformulações dos 4 modelos anteriores e aplicam-se ao dimensionamento de redes ATM que permitam reconfiguração dinâmica de VPCs e cujo tráfego esperado tenha um perfil multi-horário. Como a escolha das árvores de suporte num determinado período horário é independente das árvores de suporte escolhidas nos outras períodos horários, estas formulações são extensões das anteriores em que todas as variáveis associadas às redes lógicas e às classes de tráfego são afectadas de um índice adicional (índice h) que representa o período horário a que dizem respeito.

5.4.1 Modelo A de Dimensionamento Multi-Horário

N Conjunto de nós da rede

A Conjunto de arestas $\{i,j\}$ da rede: (i,j) representa a aresta $\{i,j\}$ no sentido de i para j e (j,i) representa o sentido inverso

T Conjunto dos tipos t de ligações físicas possíveis de implementar entre os nós de rede

α_t Largura de banda de uma ligação física do tipo t

$Y_{\{ij\}}^t$ Número máximo de ligações físicas do tipo t os nós i e j

$C_{\{ij\}}^t$ Custo associado à utilização de uma ligação do tipo t entre os nós i e j

S Conjunto de serviços ponto-multiponto s suportados pela rede

H Conjunto de períodos horários h

M_s Diâmetro máximo da rede lógica de suporte ao serviço s

S_u Conjunto dos nós de tráfego do serviço s

$\underline{S}_u$ Conjunto dos nós de trânsito do serviço s ($S_u \cup \underline{S}_u = N$ e $S_u \cap \underline{S}_u = \varnothing$)

Ω_n^s Máximo tráfego (em Erlangs) suportado por um VPC com capacidade para n ligações de suporte ao serviço s

$\underline{n}_s^h$ Número mínimo de ligações de qualquer VPC de suporte ao serviço s no período horário h

$\overline{n}_s^h$ Número máximo de ligações de qualquer VPC de suporte ao serviço s no período horário h

R_k^s Conjunto das combinações r de nós pertencentes a $S_u \setminus \{k\}$ para as quais existe tráfego oferecido no nó $k \in S_u$

ρ_k^{srh} Tráfego oferecido (em Erlangs) pelo serviço s no nó origem $k \in S_u$ para a combinação de nós destino $r \in R_k^s$ no período horário h

w_s Largura de banda de uma ligação do serviço s

$y_{\{ij\}}^t$ Variável inteira que determina o número de ligações físicas do tipo t entre os nós i e j

p_{ij}^{skh} Variável binária: o valor 1 significa que, para o serviço s no período horário h, a árvore de serviço com origem no nó $k \in S_u$ passa no arco (i,j)

x_{ij}^{skrh} Variável binária: o valor 1 significa que, para o serviço s no período horário h, a árvore de estabelecimento de ligação do nó $k \in S_u$ para a combinação de nós destino $r \in R_k^s$ passa no arco (i,j)

x_{ij}^{skdh} Variável binária: o valor 1 significa que, para o serviço s no período horário h, o trajecto do nó $k \in S_u$ para o nó $d \in S_u$ incluído nas árvores de estabelecimento de ligação das classes com origem em k e cujo conjunto de destinos r contenha d, passa no arco (i,j)

$z_{\{ij\}}^{sh}$ Variável binária de árvore de suporte: o valor 1 significa que a aresta $\{i,j\}$ pertence à árvore de suporte do serviço s no período horário h

u_{ij}^{snh} Variável binária de capacidade: o valor 1 significa que o arco (i,j) da árvore de suporte do serviço s do período horário h tem capacidade para n ligações simultâneas

Minimizar $\sum_{\{i,j\}\in A}\sum_{t\in T} C_{\{ij\}}^t y_{\{ij\}}^t$

Sujeito a:

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{kj}^{skdh} - x_{jk}^{skdh}\right) = 1 \qquad , s \in S, h \in H, k \in S_u, d \in S_u \setminus \{k\} \qquad \text{(20a)}$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{jd}^{skdh} - x_{dj}^{skdh}\right) = 1 \qquad , s \in S, h \in H, k \in S_u, d \in S_u \setminus \{k\} \qquad \text{(20b)}$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{ij}^{skdh} - x_{ji}^{skdh}\right) = 0 \qquad , s \in S, h \in H, k \in S_u, d \in S_u \setminus \{k\}, i \neq k \text{ e } i \neq d \qquad \text{(20c)}$$

$$x_{ij}^{skdh} \leq x_{ij}^{skrh} \qquad , s \in S, h \in H, k \in S_u, r \in R_k^s : d \in r, (i,j) \in A \qquad \text{(20d)}$$

$$x_{ij}^{skdh} \leq p_{ij}^{skh} \qquad , s \in S, h \in H, k \in S_u, d \in S_u \setminus \{k\}, (i,j) \in A \qquad \text{(20e)}$$

$$\sum_{j\in N:\{i,j\}\in A} p_{ji}^{skh} \leq 1 \qquad , s \in S, h \in H, k \in S_u, i \in N \qquad \text{(20f)}$$

$$\sum_{j \in N:\{i,j\} \in A} p_{ij}^{skh} \leq 1 \qquad , s \in S\,,\, h \in H\,,\, k \in S_u\,,\, i \in \underline{S}_u \qquad (20g)$$

$$p_{ij}^{skh} + p_{ji}^{skh} = z_{\{ij\}}^{sh} \qquad , s \in S\,,\, h \in H\,,\, k \in S_u\,,\, \{i,j\} \in A \qquad (20h)$$

$$\sum_{i \in S_u} \sum_{j \in N:\{i,j\} \in A} x_{ji}^{skdh} \leq M_s \qquad , s \in S\,,\, h \in H\,,\, k \in S_u\,,\, d \in S_u \setminus \{k\} \qquad (20i)$$

$$\sum_{k \in S_u} \sum_{r \in R_k^s} \rho_k^{srh} \cdot x_{ij}^{skrh} \leq \sum_{n=\underline{n}_s^h}^{\bar{n}_s^h} \Omega_n^s \cdot u_{ij}^{snh} \qquad , s \in S\,,\, h \in H\,,\, (i,j) \in A \qquad (20j)$$

$$\sum_{n=\underline{n}_s^h}^{\bar{n}_s^h} u_{ij}^{snh} \leq 1 \qquad , s \in S\,,\, h \in H\,,\, (i,j) \in A \qquad (20k)$$

$$\sum_{s \in S} \sum_{h \in H} \sum_{n=\underline{n}_s^h}^{\bar{n}_s^h} \left(w_s \cdot n \cdot u_{ij}^{snh} \right) \leq \sum_{t \in T} \alpha_t y_{\{ij\}}^t \qquad , (i,j) \in A \qquad (20l)$$

$$y_{\{ij\}}^t \leq Y_{\{ij\}}^t \qquad , \{i,j\} \in A\,,\, t \in T \qquad (20m)$$

$x_{ij}^{skrh} \in \{0,1\}$, $p_{ij}^{skh} \in \{0,1\}$, $z_{\{ij\}}^{sh} \in \{0,1\}$, $u_{ij}^{snh} \in \{0,1\}$, $y_{\{ij\}}^t$ inteiro não negativo

De notar que o trajecto e a capacidade dos VPCs que constituem cada árvore de suporte não são directamente modelizados na formulação apresentada. No entanto, estas grandezas estão indirectamente representadas pelas variáveis $z_{\{ij\}}^{sh}$ e u_{ij}^{snh}. Para cada período horário *h*, cada conjunto de arestas da árvore de suporte $\{\{i,j\}: z_{\{ij\}}^{sh}=1\}$ que definam um trajecto entre dois nós de tráfego do serviço *s*, define o trajecto de um VPC. Para cada um dos VPCs, a grandeza:

$$\sum_{n=\underline{n}_s^h}^{\bar{n}_s^h} \left(w_s \cdot n \cdot u_{ij}^{snh} \right)$$

define a capacidade do VPC no arco (*i*,*j*) por onde passa no período horário *h*.

5.4.2 Modelo B de Dimensionamento Multi-Horário

N Conjunto de nós da rede

A Conjunto de arestas {*i*,*j*} da rede: (*i*,*j*) representa a aresta {*i*,*j*} no sentido de *i* para *j* e (*j*,*i*) representa o sentido inverso

T Conjunto dos tipos *t* de ligações físicas possíveis de implementar entre os nós de rede

α_t Largura de banda de uma ligação física do tipo *t*

$Y_{\{ij\}}^t$ Número máximo de ligações físicas do tipo *t* entre os nós *i* e *j*

$C_{\{ij\}}^t$ Custo associado à utilização de uma ligação do tipo *t* entre os nós *i* e *j*

S Conjunto de serviços ponto-multiponto *s* suportados pela rede

H Conjunto de períodos horários *h*

M_s Diâmetro máximo da rede lógica de suporte ao serviço *s*

S_u Conjunto dos nós de tráfego do serviço *s*

$\underline{S}_u$ Conjunto dos nós de trânsito do serviço *s* ($S_u \cup \underline{S}_u = N$ e $S_u \cap \underline{S}_u = \varnothing$)

$\underline{n}_s^h$ Número mínimo de ligações de qualquer VPC de suporte ao serviço *s* no período horário *h*

$\Phi_{\underline{n}}^{sh}$ Máximo tráfego (em Erlangs) suportado por um VPC de suporte ao serviço *s* no período horário *h* com capacidade para $\underline{n}_s^h$ ligações simultâneas

Φ_{Δ}^{sh} Capacidade de tráfego (em Erlangs) adicional suportado por um VPC de suporte ao serviço *s* no período horário *h* para cada ligação acima do valor mínimo de $\underline{n}_s^h$ ligações

R_k^s Conjunto das combinações *r* de nós pertencentes a $S_u \setminus \{k\}$ para as quais existe tráfego oferecido no nó $k \in S_u$

ρ_k^{srh} Tráfego oferecido (em Erlangs) pelo serviço *s* no nó origem $k \in S_u$ para a combinação de nós destino $r \in R_k^s$ no período horário *h*

w_s Largura de banda de uma ligação do serviço *s*

$y_{\{ij\}}^t$ Variável inteira que determina o número de ligações físicas do tipo *t* entre os nós *i* e *j*

p_{ij}^{skh} Variável binária: o valor 1 significa que, para o serviço *s* no período horário *h*, a árvore de serviço com origem no nó $k \in S_u$ passa no arco (*i*,*j*)

x_{ij}^{skrh} Variável binária: o valor 1 significa que, para o serviço *s* no período horário *h*, a árvore de estabelecimento de ligação do nó $k \in S_u$ para a combinação de nós destino $r \in R_k^s$ passa no arco (*i*,*j*)

x_{ij}^{skdh} Variável binária: o valor 1 significa que, para o serviço *s* no período horário *h*, o trajecto do nó $k \in S_u$ para o nó $d \in S_u$ incluído nas árvores de estabelecimento de ligação das classes com origem em *k* e cujo conjunto de destinos *r* contenha *d*, passa no arco (*i*,*j*)

$z_{\{ij\}}^{sh}$ Variável binária de árvore de suporte: o valor 1 significa que a aresta $\{i,j\}$ pertence à árvore de suporte do serviço *s* no período horário *h*

u_{ij}^{sh} Variável binária de capacidade: o valor 1 significa que o arco (i,j) da árvore de suporte do serviço *s* do período horário *h* tem capacidade para pelo menos $\underline{n}_s^h$ ligações simultâneas

$u_{ij}^{s\Delta h}$ Variável real de capacidade cujo valor representa a capacidade adicional de ligações no arco (i,j) da árvore de suporte do serviço *s* do período horário *h* acima das $\underline{n}_s^h$ ligações simultâneas

Minimizar $\sum_{\{i,j\}\in A}\sum_{t\in T} C_{\{ij\}}^t y_{\{ij\}}^t$

Sujeito a:

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{kj}^{skdh} - x_{jk}^{skdh}\right)=1 \qquad , s\in S\,,\, h\in H\,,\, k\in S_u\,,\, d\in S_u\setminus\{k\} \qquad (20a)$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{jd}^{skdh} - x_{dj}^{skdh}\right)=1 \qquad , s\in S\,,\, h\in H\,,\, k\in S_u\,,\, d\in S_u\setminus\{k\} \qquad (20b)$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{ij}^{skdh} - x_{ji}^{skdh}\right)=0 \qquad , s\in S\,,\, h\in H\,,\, k\in S_u\,,\, d\in S_u\setminus\{k\}\,,\, i\neq k \text{ e } i\neq d \qquad (20c)$$

$$x_{ij}^{skdh} \le x_{ij}^{skrh} \qquad , s\in S\,,\, h\in H\,,\, k\in S_u\,,\, r\in R_k^s : d\in r\,,(i,j)\in A \qquad (20d)$$

$$x_{ij}^{skdh} \le p_{ij}^{skh} \qquad , s\in S\,,\, h\in H\,,\, k\in S_u\,,\, d\in S_u\setminus\{k\}\,,\, (i,j)\in A \qquad (20e)$$

$$\sum_{j\in N:\{i,j\}\in A} p_{ji}^{skh} \le 1 \qquad , s\in S\,,\, h\in H\,,\, k\in S_u\,,\, i\in N \qquad (20f)$$

$$\sum_{j\in N:\{i,j\}\in A} p_{ij}^{skh} \le 1 \qquad , s\in S\,,\, h\in H\,,\, k\in S_u\,,\, i\in \underline{S}_u \qquad (20g)$$

$$p_{ij}^{skh} + p_{ji}^{skh} = z_{\{ij\}}^{sh} \qquad , s\in S\,,\, h\in H\,,\, k\in S_u\,,\, \{i,j\}\in A \qquad (20h)$$

$$\sum_{i\in S_u}\sum_{j\in N:\{i,j\}\in A} x_{ji}^{skdh} \le M_s \qquad , s\in S\,,\, h\in H\,,\, k\in S_u\,,\, d\in S_u\setminus\{k\} \qquad (20i)$$

$$\sum_{k\in S_u}\sum_{r\in R_k^s} \rho_k^{srh}\cdot x_{ij}^{skrh} \le \Phi_{\underline{n}}^{sh}\cdot u_{ij}^{sh} + \Phi_{\Delta}^{sh}\cdot u_{ij}^{s\Delta h} \qquad , s\in S\,,\, h\in H\,,\, (i,j)\in A \qquad (20j')$$

$$u_{ij}^{sh} \ge 0.0001\cdot u_{ij}^{s\Delta h} \qquad , s\in S\,,\, h\in H\,,\, (i,j)\in A \qquad (20k')$$

$$\sum_{s\in S}\sum_{h\in H} w_s\left(\underline{n}_s^h \cdot u_{ij}^{sh} + u_{ij}^{s\Delta h}\right) \le \sum_{t\in T} \alpha_t y_{\{ij\}}^t \qquad , (i,j)\in A \qquad (20l')$$

$$y_{\{ij\}}^t \le Y_{\{ij\}}^t \qquad , \{i,j\}\in A\ , t\in T \qquad (20m)$$

$x_{ij}^{skrh} \in \{0,1\}$, $p_{ij}^{skh} \in \{0,1\}$, $z_{\{ij\}}^{sh} \in \{0,1\}$, $u_{ij}^{sh} \in \{0,1\}$

$u_{ij}^{s\Delta h}$ real não negativo $y_{\{ij\}}^t$ inteiro não negativo

De notar que o trajecto e a capacidade dos VPCs que constituem cada árvore de suporte não são directamente modelizados na formulação apresentada. No entanto, estas grandezas estão indirectamente representadas pelas variáveis $z_{\{ij\}}^{sh}$, u_{ij}^{sh} e $u_{ij}^{s\Delta h}$. Para cada período horário *h*, cada conjunto de arestas da árvore de suporte $\{\{i,j\}: z_{\{ij\}}^{sh}=1\}$ que definam um trajecto entre dois nós de tráfego do serviço *s*, define o trajecto de um VPC. Para cada um dos VPCs, a grandeza:

$$w_s\left(\underline{n}_s^h \cdot u_{ij}^{sh} + u_{ij}^{s\Delta h}\right)$$

define a capacidade do VPC no arco (*i*,*j*) por onde passa no período horário *h*.

5.4.3 Modelo A Reformulado de Dimensionamento Multi-Horário

N	Conjunto de nós da rede
A	Conjunto de arestas $\{i,j\}$ da rede: (i,j) representa a aresta $\{i,j\}$ no sentido de *i* para *j* e (j,i) representa o sentido inverso
T	Conjunto dos tipos *t* de ligações físicas possíveis de implementar entre os nós de rede
α_t	Largura de banda de uma ligação física do tipo *t*
$Y_{\{ij\}}^t$	Número máximo de ligações físicas do tipo *t* entre os nós *i* e *j*
$C_{\{ij\}}^t$	Custo associado à utilização de uma ligação do tipo *t* entre os nós *i* e *j*
S	Conjunto de serviços ponto-multiponto *s* suportados pela rede
H	Conjunto de períodos horários *h*
M_s	Diâmetro máximo da rede lógica de suporte ao serviço *s*
S_u	Conjunto dos nós de tráfego do serviço *s*
$\underline{S}_u$	Conjunto dos nós de trânsito do serviço *s* ($S_u \cup \underline{S}_u = N$ e $S_u \cap \underline{S}_u = \varnothing$)
$\underline{n}_s^h$	Número mínimo de ligações de qualquer VPC de suporte ao serviço *s* no período horário *h*

$\bar{n}_s^h$ Número máximo de ligações de qualquer VPC de suporte ao serviço *s* no período horário *h*

$\Omega_{\underline{n}}^{sh}$ Máximo tráfego (em Erlangs) suportado por um VPC de suporte ao serviço *s* no período horário *h* com capacidade para $\underline{n}_s^h$ ligações

$\underline{\Omega}_n^{sh}$ Capacidade de tráfego (em Erlangs) adicional suportado por um VPC de suporte ao serviço *s* no período horário *h* com capacidade para *n* ligações. Este valor é a diferença entre o tráfego suportado por *n* ligações e o tráfego suportado por $\underline{n}_s^h$ ligações

R_k^s Conjunto das combinações *r* de nós pertencentes a $S_u \setminus \{k\}$ para as quais existe tráfego oferecido no nó $k \in S_u$

ρ_k^{srh} Tráfego oferecido (em Erlangs) pelo serviço *s* no período horário *h* no nó origem $k \in S_u$ para a combinação de nós destino $r \in R_k^s$

w_s Largura de banda de uma ligação do serviço *s*

$y_{\{ij\}}^t$ Variável inteira que determina o número de ligações físicas do tipo *t* entre os nós *i* e *j*

p_{ij}^{skh} Variável binária: o valor 1 significa que, para o serviço *s* no período horário *h*, a árvore de serviço com origem no nó $k \in S_u$ passa no arco (*i*,*j*)

x_{ij}^{skrh} Variável binária: o valor 1 significa que, para o serviço *s* no período horário *h*, a árvore de estabelecimento de ligação do nó $k \in S_u$ para a combinação de nós destino $r \in R_k^s$ passa no arco (*i*,*j*)

x_{ij}^{skdh} Variável binária: o valor 1 significa que, para o serviço *s* no período horário *h*, o trajecto do nó $k \in S_u$ para o nó $d \in S_u$ incluído nas árvores de estabelecimento de ligação das classes com origem em *k* e cujo conjunto de destinos *r* contenha *d*, passa no arco (*i*,*j*)

$z_{\{ij\}}^{sh}$ Variável binária de árvore de suporte: o valor 1 significa que a aresta {*i*,*j*} pertence à árvore de suporte do serviço *s* no período horário *h*

u_{ij}^{snh} Variável binária de capacidade: o valor 1 significa que o arco (*i*,*j*) da árvore de suporte do serviço *s* do período horário *h* tem capacidade para *n* ligações simultâneas

Minimizar $$\sum_{\{i,j\} \in A} \sum_{t \in T} C_{\{ij\}}^t y_{\{ij\}}^t$$

Sujeito a:

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{kj}^{skdh}-x_{jk}^{skdh}\right)=1 \qquad ,s\in S\,,h\in H\,,k\in S_u\,,d\in S_u\setminus\{k\} \tag{21a}$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{jd}^{skdh}-x_{dj}^{skdh}\right)=1 \qquad ,s\in S\,,h\in H\,,k\in S_u\,,d\in S_u\setminus\{k\} \tag{21b}$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{ij}^{skdh}-x_{ji}^{skdh}\right)=0 \qquad ,s\in S\,,h\in H\,,k\in S_u\,,d\in S_u\setminus\{k\}\,,i\neq k \text{ e } i\neq d \tag{21c}$$

$$x_{ij}^{skdh}\le x_{ij}^{skrh} \qquad ,s\in S\,,h\in H\,,k\in S_u\,,r\in R_k^s: d\in r\,,(i,j)\in A \tag{21d}$$

$$x_{ij}^{skdh}\le p_{ij}^{skh} \qquad ,s\in S\,,h\in H\,,k\in S_u\,,d\in S_u\setminus\{k\}\,,(i,j)\in A \tag{21e}$$

$$\sum_{j\in N:\{i,j\}\in A}p_{ji}^{skh}\le 1 \qquad ,s\in S\,,h\in H\,,k\in S_u\,,i\in N \tag{21f}$$

$$\sum_{j\in N:\{i,j\}\in A}p_{ij}^{skh}\le 1 \qquad ,s\in S\,,h\in H\,,k\in S_u\,,i\in \underline{S}_u \tag{21g}$$

$$p_{ij}^{skh}+p_{ji}^{skh}=z_{\{ij\}}^{sh} \qquad ,s\in S\,,h\in H\,,k\in S_u\,,\{i,j\}\in A \tag{21h}$$

$$\sum_{i\in S_u}\sum_{j\in N:\{i,j\}\in A}x_{ji}^{skdh}\le M_s \qquad ,s\in S\,,h\in H\,,k\in S_u\,,d\in S_u\setminus\{k\} \tag{21i}$$

$$\sum_{k\in S_u}\sum_{r\in R_k^s}\rho_k^{srh}\cdot x_{ij}^{skrh}\le \Omega_{\underline{n}}^{sh}\cdot z_{\{ij\}}^{sh}+\sum_{n=\underline{n}_s^h+1}^{\bar{n}_s^h}\underline{\Omega}_n^{sh}\cdot u_{ij}^{snh} \qquad ,s\in S\,,h\in H\,,(i,j)\in A \tag{21j}$$

$$\sum_{n=\underline{n}_s^h+1}^{\bar{n}_s^h}u_{ij}^{snh}\le z_{\{ij\}}^{sh} \qquad ,s\in S\,,h\in H\,,(i,j)\in A \tag{21k}$$

$$\sum_{s\in S}\sum_{h\in H}w_s\left(\underline{n}_s^h\cdot z_{\{ij\}}^{sh}+\sum_{n=\underline{n}_s^h+1}^{\bar{n}_s^h}\left[\left(n-\underline{n}_s^h\right)\cdot u_{ij}^{snh}\right]\right)\le\sum_{t\in T}\alpha_t y_{\{ij\}}^t \qquad ,(i,j)\in A \tag{21l}$$

$$y_{\{ij\}}^t\le Y_{\{ij\}}^t \qquad ,\{i,j\}\in A\,,t\in T \tag{21m}$$

$x_{ij}^{skrh}\in\{0,1\}$, $p_{ij}^{skh}\in\{0,1\}$, $z_{\{ij\}}^{sh}\in\{0,1\}$, $u_{ij}^{snh}\in\{0,1\}$, $y_{\{ij\}}^t$ inteiro não negativo

5.4.4 Modelo B Reformulado de Dimensionamento Multi-Horário

N Conjunto de nós da rede

A Conjunto de arestas $\{i,j\}$ da rede: (i,j) representa a aresta $\{i,j\}$ no sentido de i para j e (j,i) representa o sentido inverso

T Conjunto dos tipos *t* de ligações físicas possíveis de implementar entre os nós de rede

α_t Largura de banda de uma ligação física do tipo *t*

$Y_{\{ij\}}^t$ Número máximo de ligações físicas do tipo *t* entre os nós *i* e *j*

$C_{\{ij\}}^t$ Custo associado à utilização de uma ligação do tipo *t* entre os nós *i* e *j*

S Conjunto de serviços ponto-multiponto *s* suportados pela rede

H Conjunto de períodos horários *h*

M_s Diâmetro máximo da rede lógica de suporte ao serviço *s*

S_u Conjunto dos nós de tráfego do serviço *s*

$\underline{S}_u$ Conjunto dos nós de trânsito do serviço *s* ($S_u \cup \underline{S}_u = N$ e $S_u \cap \underline{S}_u = \varnothing$)

$\underline{n}_s^h$ Número mínimo de ligações de qualquer VPC de suporte ao serviço *s* no período horário *h*

$\Phi_{\underline{n}}^{sh}$ Máximo tráfego (em Erlangs) suportado por um VPC de suporte ao serviço *s* no período horário *h* com capacidade para $\underline{n}_s^h$ ligações simultâneas

Φ_{Δ}^{sh} Capacidade de tráfego (em Erlangs) adicional suportado por um VPC de suporte ao serviço *s* no período horário *h* para cada ligação acima do valor mínimo de $\underline{n}_s^h$ ligações

R_k^s Conjunto das combinações *r* de nós pertencentes a $S_u \setminus \{k\}$ para as quais existe tráfego oferecido no nó $k \in S_u$

ρ_k^{srh} Tráfego oferecido (em Erlangs) pelo serviço *s* no período horário *h* no nó origem $k \in S_u$ para a combinação de nós destino $r \in R_k^s$

w_s Largura de banda de uma ligação do serviço *s*

$y_{\{ij\}}^t$ Variável inteira que determina o número de ligações físicas do tipo *t* entre os nós *i* e *j*

p_{ij}^{skh} Variável binária: o valor 1 significa que, para o serviço *s* no período horário *h*, a árvore de serviço com origem no nó $k \in S_u$ passa no arco (*i*,*j*)

x_{ij}^{skrh} Variável binária: o valor 1 significa que, para o serviço *s* no período horário *h*, a árvore de estabelecimento de ligação do nó $k \in S_u$ para a combinação de nós destino $r \in R_k^s$ passa no arco (*i*,*j*)

x_{ij}^{skdh} Variável binária: o valor 1 significa que, para o serviço *s* no período horário *h*, o trajecto do nó $k \in S_u$ para o nó $d \in S_u$ incluído nas árvores de estabelecimento de ligação das classes com origem em *k* e cujo conjunto de destinos *r* contenha *d*, passa no arco (*i*,*j*)

$z_{\{ij\}}^{sh}$ Variável binária de árvore de suporte: o valor 1 significa que a aresta $\{i,j\}$ pertence à árvore de suporte do serviço s no período horário h

$u_{ij}^{s\Delta h}$ Variável real de capacidade cujo valor representa a capacidade adicional de ligações no arco (i,j) da árvore de suporte do serviço s do período horário h acima das $\underline{n}_s^h$ ligações simultâneas

Minimizar $\sum_{\{i,j\}\in A}\sum_{t\in T} C_{\{ij\}}^t y_{\{ij\}}^t$

Sujeito a:

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{kj}^{skdh} - x_{jk}^{skdh}\right) = 1 \qquad , s\in S, h\in H, k\in S_u, d\in S_u\setminus\{k\} \qquad (21a)$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{jd}^{skdh} - x_{dj}^{skdh}\right) = 1 \qquad , s\in S, h\in H, k\in S_u, d\in S_u\setminus\{k\} \qquad (21b)$$

$$\sum_{j\in N:\{i,j\}\in A}\left(x_{ij}^{skdh} - x_{ji}^{skdh}\right) = 0 \qquad , s\in S, h\in H, k\in S_u, d\in S_u\setminus\{k\}, i\neq k \text{ e } i\neq d \qquad (21c)$$

$$x_{ij}^{skdh} \le x_{ij}^{skrh} \qquad , s\in S, h\in H, k\in S_u, r\in R_k^s : d\in r, (i,j)\in A \qquad (21d)$$

$$x_{ij}^{skdh} \le p_{ij}^{skh} \qquad , s\in S, h\in H, k\in S_u, d\in S_u\setminus\{k\}, (i,j)\in A \qquad (21e)$$

$$\sum_{j\in N:\{i,j\}\in A} p_{ji}^{skh} \le 1 \qquad , s\in S, h\in H, k\in S_u, i\in N \qquad (21f)$$

$$\sum_{j\in N:\{i,j\}\in A} p_{ij}^{skh} \le 1 \qquad , s\in S, h\in H, k\in S_u, i\in \underline{S}_u \qquad (21g)$$

$$p_{ij}^{skh} + p_{ji}^{skh} = z_{\{ij\}}^{sh} \qquad , s\in S, h\in H, k\in S_u, \{i,j\}\in A \qquad (21h)$$

$$\sum_{i\in S_u}\sum_{j\in N:\{i,j\}\in A} x_{ji}^{skdh} \le M_s \qquad , s\in S, h\in H, k\in S_u, d\in S_u\setminus\{k\} \qquad (21i)$$

$$\sum_{k\in S_u}\sum_{r\in R_k^s} \rho_k^{srh}\cdot x_{ij}^{skrh} \le \Phi_{\underline{n}}^{sh}\cdot z_{\{ij\}}^{sh} + \Phi_{\Delta}^{sh}\cdot u_{ij}^{s\Delta h} \qquad , s\in S, h\in H, (i,j)\in A \qquad (21j')$$

$$\sum_{s\in S}\sum_{h\in H} w_s\left(\underline{n}_s^h\cdot z_{\{ij\}}^{sh} + u_{ij}^{s\Delta h}\right) \le \sum_{t\in T}\alpha_t y_{\{ij\}}^t \qquad , (i,j)\in A \qquad (21l')$$

$$y_{\{ij\}}^t \le Y_{\{ij\}}^t \qquad , \{i,j\}\in A, t\in T \qquad (21m)$$

$x_{ij}^{skrh} \in \{0,1\}$, $p_{ij}^{skh} \in \{0,1\}$, $z_{\{ij\}}^{sh} \in \{0,1\}$

$u_{ij}^{s\Delta h}$ real não negativo $y_{\{ij\}}^t$ inteiro não negativo

5.5 Resultados Computacionais

Com o objectivo de comparar o desempenho dos diferentes modelos propostos, foi dimensionada uma rede com 5 nós e 6 arestas com a topologia apresentada na Figura 5-7 com possibilidade de utilização de ligações físicas PDH E3 de 34 Mbps entre os nós da rede e sem restrições no número máximo de ligações. Os custos $C^t_{\{ij\}}$ foram calculados atribuindo um custo de comutação de 3 para todos os nós e considerando ligações dedicadas com um custo de 0.2 por unidade de comprimento (os valores de comprimento são apresentados na Figura 5-7).

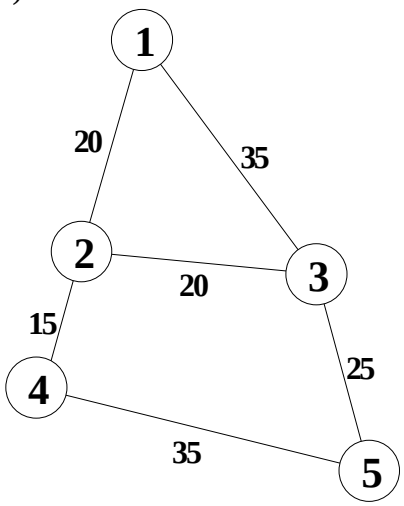


Figura 5-7 : Rede de teste com 5 nós

Consideramos 1 período horário e 2 serviços ponto-multiponto, cada um com 4 nós de tráfego e uma largura de banda por ligação lógica de 2 Mbps. Consideramos a probabilidade máxima de bloqueio de 1% para ambos os serviços e, dada a dimensão da rede, não imposemos limite ao diâmetro máximo das redes lógicas de suporte.

Foram gerados 10 problemas de dimensionamento. Em cada problema, foram seleccionados aleatoriamente 4 nós de tráfego para cada serviço e foram gerados cenários de tráfego segundo o método proposto no capítulo 2 (secção 2.4.2) em que se considerou:

- Geração de valores de tráfego de cada origem para cada destino individual dado por uma função densidade de probabilidade uniforme entre 2 e 10 Erlangs para os dois serviços.
- Geração de valores de tráfego para múltiplos destinos com os parâmetros: número máximo de destinos = 2 ; número médio de destinos = 1.5.

A Tabela 5-1 apresenta todas as grandezas relevantes que se retiram do cálculo de cada um dos 10 problemas gerados em cada um dos modelos propostos, nomeadamente: o valor da relaxação linear do problema (RL), o valor da solução óptima do problema (O) e o tempo de cálculo (TC)

	VPCs Ext-a-Ext		Modelo A	Modelo A Reformulado	Modelo B	Modelo B Reformulado
1	395	RL: O: TC:	312.32 329 884 s	320.59 329 3612 s	288.87 352 80 s	326.17 352 52 s
2	319	RL: O: TC:	244.52 276 7034 s	254.88 276 900 s	224.78 279 169 s	260.76 279 23 s
3	385	RL: O: TC:	325.37 377 34896 s	346.76 377 5644 s	303.49 381 233 s	354.90 381 62 s
4	333	RL: O: TC:	243.30 276 26190 s	253.54 276 826 s	227.68 276 126 s	260.28 276 12 s
5	349	RL: O: TC:	266.32 316 152638 s	276.03 316 4980 s	245.59 326 470 s	281.86 326 59 s
6	317	RL: O: TC:	217.86 249 16092 s	229.45 249 572 s	202.69 249 93 s	235.85 249 26 s
7	327	RL: O: TC:	246.97 290 69956 s	258.49 290 33602 s	230.46 290 216 s	265.34 290 67 s
8	374	RL: O: TC:	292.18 331 2358 s	309.60 331 4118 s	269.51 331 225 s	316.91 331 44 s
9	383	RL: O: TC:	291.96 330-343* 236000 s	307.05 343 1930 s	271.12 343 107 s	314.17 343 88 s
10	321	RL: O: TC:	249.00 286 19180 s	259.47 286 991 s	231.19 295 123 s	265.48 295 85 s

* *Ao fim de 236000 seg., o algoritmo de branch-and-bound determinou apenas estes valores como limite inferior e limite superior*

Tabela 5-1 : Relaxação linear (RL), solução óptima (O) e tempo de cálculo (TC) dos modelos para redes lógicas em árvore e solução óptima com redes lógicas de VPCs extremo-a-extremo (rede de 5 nós)

A Tabela 5-1 apresenta também os valores das soluções óptimas para o dimensionamento da rede assumindo redes lógicas com VPCs extremo-a-extremo. Neste caso, foi usada a formulação por modelização explícita de VPCs apresentada no capítulo 3.

Os resultados foram obtidos através da ferramenta CPLEX usando o algoritmo de *branch-and-bound* a correr num PC a 200 MHz com 128 Mbytes de RAM e o sistema operativo Windows. As principais conclusões que se retiram dos resultados apresentados na Tabela 5-1 são:

- As soluções dos modelos baseados na aproximação real à função ErlangB inversa (modelos B) são sempre de custo maior ou igual que os custos das soluções dos modelos baseados na aproximação binária (modelos A). Esta conclusão reflecte o sobre-dimensionamento adicional que o modelo real introduz no cálculo das capacidades dos VPCs comparativamente com o modelo binário.
- Os custos das soluções dos modelos reformulados são iguais aos custos das soluções dos seus modelos naturais. Esta conclusão reforça a validade das reformulações propostas.
- Comparando o modelo A natural e o modelo B natural, verificamos que embora o valor da relaxação linear de B seja substancialmente menor, este modelo exige muito menos tempo de cálculo. Este resultado é consequência da menor complexidade do modelo B devido ao menor número de variáveis binários nele contido.
- Em ambos os modelos reformulados, o valor da relaxação linear aumentou em comparação com os seus modelos naturais e, como consequência dado que não houve aumento de complexidade, os tempos de cálculo diminuiram bastante. Este resultado significa que as reformulações propostas de facto reduziram significativamente o espaço de soluções possíveis da relaxação linear. Em especial, o modelo B reformulado mostra-se extremamente eficaz porque faz subir o valor da relaxação linear do modelo B natural (que é menor do que o do modelo A natural) para um valor que é maior do que o do modelo A reformulado.

A Tabela 5-2 faz uma comparação dos custos das soluções óptimas do dimensionamento da rede (apresentados na Tabela 5-1) baseado em redes lógicas com VPCs extremo-a-extremo e baseado nos 2 modelos reformulados propostos para redes lógicas em árvore. Estes resultados mostram que a adopção de redes lógicas em árvore permite ganhos substanciais no custo da rede dimensionada. Embora em metade dos casos o modelo B obtenha uma solução com o mesmo custo do modelo A, o modelo B é em média inferior pois obtém ganhos médios de 11.2% contra 12.6% do modelo A.

Consideramos, no entanto, que o modelo B é melhor do que o modelo A porque se obtêm soluções próximas com tempos de cálculo bastante inferiores.

	VPCs Extremo-a-Extremo	**Modelo A**	**Modelo B**
1	395	329 (16.7%)	352 (10.9%)
2	319	276 (13.5%)	279 (12.5%)
3	385	377 (2.1%)	381 (1.0%)
4	333	276 (17.1%)	276 (17.1%)
5	349	316 (9.5%)	326 (6.6%)
6	317	249 (21.5%)	249 (21.5%)
7	327	290 (11.3%)	290 (11.3%)
8	374	331 (11.5%)	331 (11.5%)
9	383	343 (11.7%)	343 (11.7%)
10	321	286 (10.9%)	295 (8.1%)
Ganho médio:		12.6%	11.2%

Tabela 5-2 : Custo das soluções óptimas dos 10 problemas gerados para a rede de 5 nós

De seguida, usamos o modelo B reformulado para dimensionar as redes de 6 nós apresentadas na Figura 5-8. O objectivo é verificar a influência do número de arestas da rede na comparação dos resultados de dimensionamento entre as redes lógicas em árvore e com VPCs extremo-a-extremo. Tal como anteriormente, consideramos a possibilidade de utilização de ligações físicas PDH E3 de 34 Mbps entre os nós da rede e sem restrições no número máximo de ligações. Os custos $C^t_{\{ij\}}$ foram calculados atribuíndo um custo de comutação de 3 para todos os nós e considerando ligações dedicadas com um custo de 0.2 por unidade de comprimento (os valores de comprimento são apresentados na Figura 5-8).

Consideramos 1 período horário e 2 serviços ponto-multiponto, um com 4 nós de tráfego e outro com 5 nós de tráfego. Consideramos uma largura de banda por ligação lógica de 2 Mbps e uma probabilidade máxima de bloqueio de 1% para ambos os serviços. Não imposemos limite máximo ao diâmetro das rede lógicas de suporte.

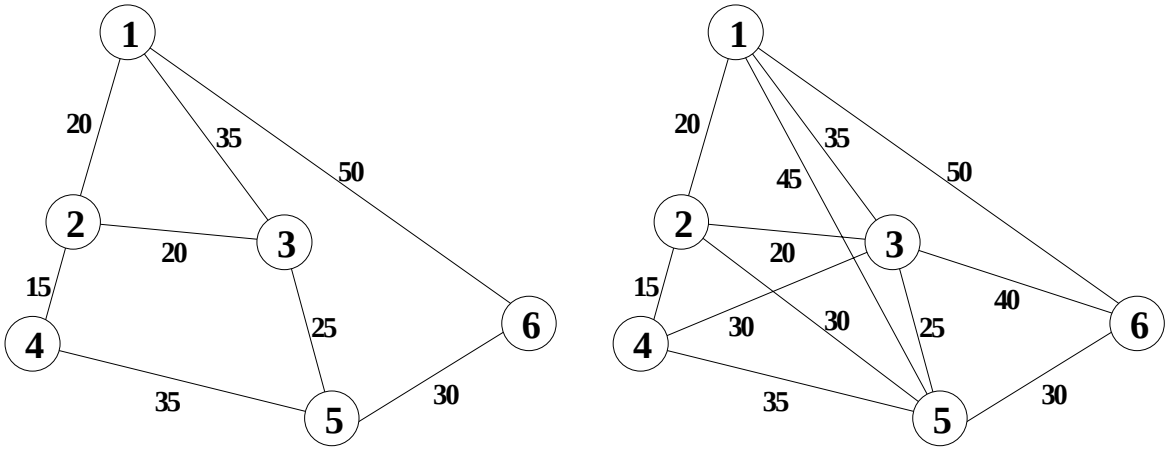


Figura 5-8 : Redes de teste com (6 nós, 8 arestas) e (6 nós, 12 arestas)

Foram gerados quatro conjuntos de 10 problemas de dimensionamento. Em todos os problemas de todos os conjuntos, foram seleccionados aleatoriamente os nós de tráfego de cada serviço para cada problema e foram gerados cenários de tráfego segundo o método proposto no capítulo 2 (secção 2.4.2). Em todos os conjuntos considerou-se um número máximo de 2 destinos para as classes de tráfego de ambos os serviços. As diferenças entre os quatro conjuntos de problemas são as que a seguir se descrevem.

O conjunto 1 de problemas considera (i) a rede de (6 nós, 8 arestas), (ii) a geração de valores de tráfego de cada origem para cada destino individual segundo uma função densidade de probabilidade uniforme entre 2 e 10 Erlangs para os dois serviços e (iii) um número médio de destinos de 1,(3) para os dois serviços.

O conjunto 2 de problemas considera (i) a rede de (6 nós, 12 arestas), (ii) a geração de valores de tráfego de cada origem para cada destino individual segundo uma função densidade de probabilidade uniforme entre 2 e 10 Erlangs para os dois serviços e (iii) um número médio de destinos de 1,(3) para os dois serviços.

O conjunto 3 de problemas considera (i) a rede de (6 nós, 8 arestas), (ii) a geração de valores de tráfego de cada origem para cada destino individual segundo uma função densidade de probabilidade uniforme entre 0.8 e 4 Erlangs para os dois serviços e (iii) um número médio de destinos de 1,(6) para os dois serviços.

O conjunto 4 de problemas considera (i) a rede de (6 nós, 12 arestas), (ii) a geração de valores de tráfego de cada origem para cada destino individual segundo uma função densidade de probabilidade uniforme entre 0.8 e 4 Erlangs para os dois serviços e (iii) um número médio de destinos de 1,(6) para os dois serviços.

Os conjuntos de problema com um número médio de 1.(6) destinos (conjuntos 3 e 4) têm limites menores para a função densidade de probabilidade uniforme porque pretendeu-se que em todos os problemas as redes lógicas com VPCs extremo-a-extremo tenham os mesmos limites inferior e superior de tráfego total suportado por cada VPC. O objectivo é quantificar a influência nos resultados do dimensionamento quando um mesmo cenário de tráfego representa, para as redes lógicas com VPCs extremo-a-extremo, perfis de tráfego ponto-multiponto com diferentes graus de diversidade. Pelo método proposto no capítulo 2 (secção 2.4.2), o valor médio do tráfego total para *n* destinos (t_n) submetido à rede para um serviço ponto-multiponto é dado por:

$$t_n = \frac{|J| \cdot C_n^m \cdot n \cdot \overline{\rho}}{n^e \cdot p^{n-1}} \qquad m = |J| - 1$$

em que $|J|$ é o número de nós de serviço, $\overline{\rho}$ é o valor médio do tráfego para um único destino e *e* e *p* são parâmetros calculados em função do número médio e número máximo de destinos para o cenário de tráfego escolhido. Como no nosso caso *n* toma apenas os valores de 1 e 2, então temos:

$$t_1 = |J|(|J|-1)\cdot\overline{\rho}$$

$$t_2 = \frac{|J|(|J|-1)(|J|-2)}{2^e \cdot p}\cdot\overline{\rho}$$

No caso dos conjuntos de problemas 1 e 2, o valor médio de destinos é 1,(3) pelo que a seguinte igualdade é válida:

$$\frac{t_1 + 2\cdot t_2}{t_1 + t_2} = \frac{4}{3} \qquad \Rightarrow \qquad 2^e \cdot p = 2\cdot(|J|-2)$$

Note-se que em redes lógicas com VPCs extremo-a-extremo, um VPC no sentido do nó *i* para o nó *j* suporta o tráfego de *i* para *j* mais o tráfego de *i* para todas as combinações de 2 nós em que um deles seja o nó *j*. Assim, o tráfego médio suportado por um VPC em qualquer um dos seus sentidos é dado por:

$$\overline{\rho} + C_1^{|J|-2}\cdot\frac{2\cdot\overline{\rho}}{2^e\cdot p} = \overline{\rho} + (|J|-2)\cdot\frac{2\cdot\overline{\rho}}{2\cdot(|J|-2)} = 2\cdot\overline{\rho}$$

Assim, nos conjuntos de problemas 1 e 2, os limites da função densidade de probabilidade uniforme entre 2 e 10 significam que o tráfego total que passa em cada sentido de cada VPC varia entre o valor mínimo de 4 e máximo de 20. Como queremos

que para os conjuntos de problemas 3 e 4 estes limites sejam os mesmos, temos de os calcular. Consideremos para os conjuntos 3 e 4 o tráfego médio de uma origem para um único destino como sendo $\overline{\rho}'$. Neste caso, o valor médio de destinos é 1,(6):

$$\frac{t_1 + 2 \cdot t_2}{t_1 + t_2} = \frac{5}{3} \qquad \Rightarrow \qquad 2^e \cdot p = \frac{(|J| - 2)}{2}$$

e o tráfego médio suportado por um VPC em qualquer um dos seus sentidos é dado por:

$$\overline{\rho}' + C_1^{|J|-2} \cdot \frac{2 \cdot \overline{\rho}'}{2^e \cdot p} = \overline{\rho}' + (|J| - 2) \cdot \frac{2 \cdot \overline{\rho}'}{(|J| - 2)/2} = 5 \cdot \overline{\rho}'$$

Como queremos que este valor seja igual em todos os conjuntos de problemas, temos que:

$$5 \cdot \overline{\rho}' = 2 \cdot \overline{\rho} \qquad \Rightarrow \qquad \overline{\rho}' = \frac{2}{5} \cdot \overline{\rho}$$

Deste modo, para que os valores mínimo e máximo do tráfego que atravessa cada VPC (numa rede lógica com VPCs extremo-a-extremo) sejam iguais em todos os conjuntos de problemas, é preciso que os limites da função densidade de probabilidade uniforme para os conjuntos 3 e 4 sejam 2/5 dos mesmos valores para os conjuntos 1 e 2. Assim resulta que estes valores nos conjuntos 3 e 4 tenham de ser 0.8 e 4, conforme foi anteriormente especificado.

As tabelas seguintes apresentam os resultados computacionais para os problemas gerados de cada conjunto. Cada tabela indica, para cada problema, o valor da solução óptima para redes lógicas com VPCs extremo-a-extremo, o valor da solução óptima para as redes lógicas em árvore, o ganho desta última em relação à primeira e o tempo de cálculo da resolução desta última. Na última linha de cada tabela são indicados os valores médios dos 10 problemas. Tal como no caso anterior, os resultados foram obtidos através da ferramenta CPLEX usando o algoritmo de *branch-and-bound* a correr num PC a 200 MHz com 128 Mbytes de RAM.

	VPCs Ext-a-Ext	Redes Lógicas Em Árvore	Tempo de Cálculo
1	390	341 (12.6%)	266 s
2	342	305 (10.8%)	296 s
3	357	307 (14.0%)	209 s
4	403	385 (4.4%)	686 s

5	378	348 (7.9%)	298 s
6	350	313 (10.6%)	57 s
7	407	362 (11.1%)	433 s
8	372	321 (13.7%)	188 s
9	401	354 (11.7%)	404 s
10	415	370 (10.8%)	352 s
Média:	381.5	340.6 (10.7 %)	319 s

Tabela 5-3 : Custo das soluções óptimas para o conjunto 1 de problemas: rede (6,8) no cenário de tráfego com número médio de destinos de 1,(3)

	VPCs Ext-a-Ext	Redes Lógicas Em Árvore	Tempo de Cálculo
1	357	350 (2.0%)	5224 s
2	361	351 (2.8%)	8016 s
3	354	329 (7.1%)	1642 s
4	374	352 (5.9%)	3611 s
5	369	347 (6.0%)	2222 s
6	354	343 (3.1%)	4439 s
7	366	315 (13.9%)	4189 s
8	337	312 (7.4%)	2471 s
9	353	320 (9.3%)	4738 s
10	317	303 (4.4%)	5253 s
Média:	354.2	332.2 (6.2 %)	4181 s

Tabela 5-4 : Custo das soluções óptimas para o conjunto 2 de problemas: rede (6,12) no cenário de tráfego com número médio de destinos de 1,(3)

	VPCs Ext-a-Ext	Redes Lógicas Em Árvore	Tempo de Cálculo
1	397	338 (14.9%)	277 s
2	417	366 (12.2%)	603 s
3	322	289 (10.2%)	391 s
4	388	352 (9.3%)	1033 s
5	364	329 (9.6%)	550 s
6	405	338 (16.5%)	377 s
7	343	298 (13.1%)	445 s
8	404	330 (18.3%)	256 s
9	393	342 (13.0%)	627 s
10	354	323 (8.8%)	1093 s
Média:	378.7	330.5 (12.7 %)	565 s

Tabela 5-5 : Custo das soluções óptimas para o conjunto 3 de problemas: rede (6,8) no cenário de tráfego com número médio de destinos de 1,(6)

	VPCs Ext-a-Ext	Redes Lógicas Em Árvore	Tempo de Cálculo
1	335	341 (-1.8%)	19361 s
2	337	297 (11.9%)	4756 s
3	346	294 (15.0%)	6324 s
4	354	315 (11.0%)	5088 s
5	373	340 (8.8%)	4949 s
6	347	316 (8.9%)	2767 s
7	359	314 (12.5%)	1406 s
8	366	336 (8.2%)	8864 s
9	348	299 (14.1%)	3966 s
10	367	353 (3.8%)	7717 s
Média:	353.2	320.5 (9.3 %)	6520 s

Tabela 5-6 : Custo das soluções óptimas para o conjunto 4 de problemas: rede (6,12) no cenário de tráfego com número médio de destinos de 1,(6)

A principal conclusão que se retira da análise dos resultados é que a adopção de redes lógicas em árvore conduz a ganhos significativos nos custos da rede física de suporte. Os resultados mostram que os ganhos resultantes da adopção de redes lógicas de suporte em árvore compensam largamente o efeito da modularidade das ligações físicas para quase todos os problemas considerados. Dos 40 problemas considerados, apenas 1 deu um resultado pior para o caso das redes lógicas em árvore. Os ganhos médios variam entre 6.2% e 12.6 % conforme os casos considerados. Este facto é ainda mais significativo se atendermos ao facto que o grau de diversidade do tráfego considerado é bastante pequeno. Na prática, é provável que um serviço ponto-multiponto tenha graus de diversidade bastante mais elevados do que os considerados nos casos apresentados.

A comparação dos resultados entre os diferentes conjuntos de problemas confirma alguns aspectos esperados antecipadamente. Os ganhos das soluções com redes lógicas em árvore são maiores para os cenários de tráfego com maior diversidade. Este resultado confirma que quanto maior for a percentagem de tráfego para múltiplos destino, mais eficiente é a adopção de redes lógicas com uma topologia em árvore.

Os ganhos das topologias em árvore são maiores para as redes mais esparsas. De facto, uma topologia em árvore tende a usar poucas arestas pois ela reduz o número de VPCs ao valor mínimo possível. Daí que a existência de mais arestas influencie menos o custo da solução óptima. Pelo contrário, uma topologia com VPCs extremo-a-extremo gera o número máximo de VPCs possível de interligação entre os nós de tráfego. Assim, a

existência de mais arestas faz com que haja muito mais trajectos possíveis para cada VPC e o efeito da modularidade das ligações físicas é mais atenuado. Os resultados mostram que comparando os resultados do dimensionamento das redes esparsas para as redes densas, a média dos custos das soluções óptimas diminui bastante mais nas redes lógicas com VPCs extremo-a-extremo do que nas redes lógicas em árvore. Este facto indicia que para redes densas com maior número de nós, é possível que a solução com VPCs extremo-a-extremo permita redes de suporte de menor custo do que com redes lógicas em árvore, pelo menos para cenários de tráfego com pequeno grau de diversidade.

Finalmente, os tempos de cálculo das redes com 6 nós (entre 5 minutos e quase 2 horas) são bem maiores que os tempos de cálculo do modelo B reformulado para o caso da rede com 5 nós (menos de 1 minuto). Isto significa que estes exemplos estão sensivelmente no limite do tamanho máximo dos problemas de dimensionamento que é possível calcular até à optimalidade com um algoritmo de *branch-and-bound*. Por esta razão, os resultados computacionais não incluem exemplos para cenários multi-horário nem consideram cenários de tráfego com mais do que 2 nós destino.

5.6 Conclusões

Como alternativa à arquitectura baseada em VPCs extremo-a-extremo apresentada no capítulo 3, neste capítulo foi proposta uma arquitectura em árvore para as redes de suporte a serviços ponto-multiponto. Neste caso, já não é possível separar o cálculo da capacidade dos VPC do dimensionamento da rede física que simultaneamente determina o trajecto dos VPCs. No entanto, foi possível determinar um conjunto de modelos de Programação Inteira ou Programação Inteira Mista capazes de incluir as restrições de qualidade de serviço ao nível de ligação (nomeadamente a probabilidade de bloqueio e o tempo de estabelecimento de ligação) na formulação do problema de dimensionamento que simultaneamente calcula as capacidades e os trajectos de cada VPC.

Com os modelos propostos, foi possível calcular as soluções óptimas para alguns casos com redes até 6 nós e foi possível demonstrar que a adopção de uma topologia em árvore para o suporte de serviços ponto-multiponto permite uma utilização mais eficiente dos recursos da rede do que uma topologia com VPCs extremo-a-extremo.

CAPÍTULO 6

CONSIDERAÇÕES FINAIS

Neste capítulo são apresentadas as principais conclusões resultantes do trabalho efectuado e são apontados os tópicos deixados em aberto e considerados mais importantes para trabalho futuro.

6.1 Principais Conclusões

Esta tese abordou o problema do dimensionamento de redes ATM num cenário multi-serviço de suporte a serviços com diferentes características de tráfego: serviços conversacionais, serviços de consulta e serviços ponto-multiponto. O problema estudado considerou a perspectiva de um operador de telecomunicações que tem como objectivo a determinação da rede física com o menor custo possível de operação e manutenção e que seja capaz de suportar o tráfego esperado para os diferentes serviços com a qualidade de serviço necessária. Os problemas de dimensionamento abordados consideraram serviços comutados em que cada serviço é suportado por uma rede lógica constituída por VPCs do tipo DBR (*Deterministic Bit Rate*) e o encaminhamento é fixo (uni-horário ou multi-horário).

O primeiro problema endereçado foi o dimensionamento de redes ATM em que as redes lógicas de suporte a todos os tipos de serviços são constituídas exclusivamente por

VPCs extremo-a-extremo. Foram propostos dois modelos de Programação Inteira para este problema. Um dos modelos baseou-se na modelização explícita dos trajectos de cada VPC. O outro modelo baseou-se na modelização dos VPCs por trajectos candidatos. Os modelos desenvolvidos extendem modelos previamente conhecidos apenas para serviços conversacionais de modo a poder ser considerado tráfego assimétrico (característico dos serviços de consulta e serviços ponto-multiponto) e a modelização de VPCs em que um dos extremos não é conhecido à partida (característico dos serviços de consulta quando existem múltiplos servidores na rede). Explorou-se a técnica da Relaxação Lagrangeana com Optimização por Sub-Gradiente para a resolução dos modelos desenvolvidos.

Os resultados computacionais mostraram que a resolução do problema de dimensionamento tem melhor desempenho no modelo baseado em modelização explícita de VPCs pois consegue em média determinar soluções de custo menor com tempos de cálculo também menores. Foi também possível ilustrar a utilidade de se considerarem modelos de dimensionamento multi-horário pelos ganhos que se obtêm quando a rede permite reconfiguração dinâmica de VPCs.

A heurística proposta contempla a determinação de um limite inferior teórico para o custo da solução óptima do problema. Usando como critério o intervalo dual, é possível ter uma medida da confiança na qualidade das soluções encontradas. Os resultados computacionais obtidos mostraram que o algoritmo proposto é eficiente na resolução de problemas de dimensionamento (tempos de cálculo pequenos) e a sua eficácia (grau de confiança na qualidade das soluções) depende principalmente da capacidade média por aresta dos VPCs a suportar na rede. Para problemas com intervalos duais grandes, não é possível saber a qualidade da solução encontrada a não ser que por outros métodos se consigam encontrar limites inferiores melhores dos que são cálculados pelo método proposto. Para as redes pequenas, foi possível através de um método alternativo, determinar limites inferiores teóricos bastante acima dos limites determinados pelo algoritmo proposto indiciando assim que a eficácia do algoritmo proposto é bastante maior do que os intervalos duais fazem parecer.

Em termos operacionais, uma arquitectura baseada apenas em VPCs extremo-a-extremo tem vantagens e desvantagens. A principal vantagem é a simplificação dos mecanismos de CAC (*Call Admission Control*) e a consequente minimização dos tempos de estabelecimento de ligação. A principal desvantagem é não ser escalável para redes de

maiores dimensões devido ao número de VPCs que é necessário configurar na rede. Em termos de dimensionamento, esta arquitectura tem também vantagens e desvantagens. A principal vantagem é a possibilidade de separar o cálculo das capacidades dos VPCs que compõem cada rede lógica do problema de dimensionamento de rede física. As principais desvantagens são não explorar ganhos de partilha de recursos que se podem obter quando diferentes classes de tráfego são suportados por VPCs comuns e, no caso de serviços ponto-multiponto, não explorar ganhos que se podem obter quando se usa a capacidade de replicação de células que os comutadores ATM permitem.

Para obviar as desvantagens duma arquitectura com VPCs extremo-a-extremo, foram propostas duas arquitecturas alternativas: uma arquitectura hierarquizada a dois níveis para redes lógicas de suporte a serviços conversacionais e uma arquitectura em árvore para redes lógicas de suporte a serviços ponto-multiponto. Para ambas as arquitecturas propostas, foram estudados modelos de optimização linear adequados que definem o problema de dimensionamento da rede.

Nenhuma das propostas permite separar o cálculo das capacidades dos VPCs do problema de dimensionamento da rede física. Na arquitectura hierarquizada a dois níveis foi possível combinar, com pequenas aproximações, estes dois problemas. Na arquitectura em árvore foram propostas aproximações que permitiram modelizar as classes de tráfego e incluir no modelo de optimização restrições que contemplem os parâmetros de Qualidade de Serviço ao nível da ligação.

Um aspecto importante no problema de dimensionamento abordado é a modularidade das ligações físicas entre nós de comutação. As capacidades das ligações físicas têm valores modulares e, no caso geral, não é possível ter uma rede física em que as capacidades das ligações físicas estejam completamente ocupadas pelos VPCs que por elas passam. Há sempre alguma capacidade nas ligações físicas que não é usada pelos VPCs. Uma rede lógica com VPCs extremo-a-extremo divide a capacidade necessária para suportar o tráfego de cada serviço num grande número de VPCs que, no caso geral, permite soluções com pequenas capacidades não usadas nas ligações físicas. Qualquer alternativa a esta arquitectura tem um número de VPCs menor e, se a capacidade média por VPC resultante for maior, é inevitavelmente menos eficiente o seu condicionamento nas ligações físicas.

No caso de redes lógicas hierarquizadas a dois níveis para suporte a serviços conversacionais, os modelos propostos foram resolvidos pelo algoritmo *branch-and-bound* para problemas uni-horário com dois serviços e redes com 10 nós. Os resultados computacionais mostraram que mesmo para redes pequenas, é possível ter ganhos de dimensionamento comparativamente à arquitectura com VPCs extremo-a-extremo. Verificou-se assim que os ganhos de partilha de recursos obtidos por esta arquitectura são maiores do que o que se perde pelas aproximações feitas na determinação do modelo de optimização e pelo efeito da modularidade das ligações físicas.

No caso de redes lógicas em árvore para suporte a serviços ponto-multiponto, os modelos propostos foram resolvidos pelo algoritmo *branch-and-bound* para problemas uni-horário com dois serviços e redes com 5 e 6 nós. Para ser possível comparar os resultados de dimensionamento nas duas arquitecturas estudadas para serviços ponto-multiponto, foi proposto um método de geração de cenários de tráfego em que é possível escolher diferentes graus de diversidade de tráfego. Pela complexidade dos modelos propostos, os cenários de tráfego gerados comtemplaram apenas classes de tráfego para um máximo de 2 destinos. Os resultados computacionais mostraram também que esta arquitectura permite ter ganhos de dimensionamento significativos comparativamente à arquitectura com VPCs extremo-a-extremo mesmo para redes pequenas e para tráfego com um grau de diversidade também bastante pequeno. Esta conclusão é mais significativa quando comparada com as redes lógicas hierarquizadas a dois níveis porque as aproximações usadas para determinar os modelos de dimensionamento são bem mais conservativas no caso das redes lógicas em árvore. Verificou-se assim que os ganhos de partilha de recursos e de utilização da replicação de células da tecnologia ATM são bastante maiores do que o que se perde pelas aproximações feitas na determinação do modelo de optimização e no efeito da modularidade das ligações físicas.

Finalmente, embora os problemas de dimensionamento tenham sido estudados seperadamente para cada arquitectura proposta, a estrutura dos modelos de programação respectivos permite que as diferentes arquitecturas possam ser combinadas num único problema de optimização. Com base nas restrições definidas para cada arquitectura, é possível enunciar o problema de dimensionamento multi-serviço em que se adoptam diferentes arquitecturas para as redes lógicas de suporte aos diferentes tipos de serviços. Por exemplo, o dimensionamento uni-horário de uma rede ATM em que o operador

pretenda adoptar uma arquitectura hierárquica para os serviços conversacionais, uma arquitectura com VPCs extremo-a-extremo para os serviços de consulta e uma arquitectura em árvore para os serviços ponto-multiponto é enunciado por um problema de dimensionamento cujo modelo de programação é dado pelas restrições (6b-k) do capítulo 4 (secção 4.4) para os VPCs pertencentes aos serviços conversacionais, pelas restrições (1b-d) do capítulo 3 (secção 3.2.1) para os VPCs pertencentes aos serviços de consulta e pelas restrições (19a-j') (secção 5.3.2) do capítulo 5 para os VPCs pertencentes aos serviços ponto-multiponto acrescentando ainda uma restrição de capacidade que garanta que a capacidade das ligações físicas em cada aresta é suficiente para suportar todos os VPCs de todas as redes lógicas que por ela passam.

6.2 Sugestões para Trabalho Futuro

Em relação ao problema de dimensionamento com redes lógicas constituídas por VPCs extremo-a-extremo é conhecido [Magnanti95] que os limites inferiores dados pelo algoritmo proposto são maus. Os resultados obtidos para redes pequenas assim o indiciam. Assim, é importante prosseguir a investigação de métodos mais eficientes no cálculo de limites inferiores do problema. Só desta forma é possível determinar se, quando os intervalos duais são grandes, as soluções obtidas pelo algoritmo proposto são próximas da solução óptima ou se vale a pena prosseguir na investigação de métodos mais eficazes de resolução do problema de dimensionamento.

Em relação aos problemas de dimensionamento com redes lógicas hierarquizadas a dois níveis e com redes lógicas em árvore, é obviamente importante investigar métodos que permitam resolver os modelos propostos para redes maiores do que as que foi possível fazer com o algoritmo de *branch-and-bound*. No primeiro caso, a Relaxação Lagrangeana aplicada às restrições de capacidade com Optimização por Sub-Gradiente parece ser uma escolha promissora se for possível determinar um critério objectivo que limite em cada região o conjunto de possíveis nós fronteira. No segundo caso, o problema é mais complexo e para ser possível avaliar a melhor forma de o resolver, é necessário ainda mais investigação em novas formas de reformulação do problema que permitam melhorar o valor da sua relaxação linear.

Ainda no caso de redes lógicas hierarquizadas a dois níveis, é importante também investigar critérios adequados para a determinação do número de regiões e dos nós de cada

região. Não é possível incluir a escolha das regiões no modelo de dimensionamento proposto porque as aproximações que serviram de base para a sua determinação deixariam de ser válidas. O estudo destes critérios terá necessariamente de contemplar a análise das matrizes de tráfego esperado de modo a determinar regiões que potencialmente resultem em maiores ganhos de partilha de recursos nos VPCs inter-regiões e nos VPCs de saída de cada região.

Em relação a outras tecnologias de telecomunicações, os modelos propostos, embora aplicados ao dimensionamento de redes ATM, parecem ser facilmente generalizáveis para qualquer rede de circuitos virtuais. No domínio das redes IP, a introdução do protocolo *Resource Reservation Protocol* (RSVP) [Zhang93] [IETF97] permite que se possam reservar recursos da rede para fluxos de informação ponto-a-ponto e ponto-multiponto. Mais recentemente, com a introdução da tecnologia *Multiprotocol Label Switching* (MPLS) [Xiao99] [IETF99] [Comm99], é possível configurar percursos virtuais (denominados *Label Switched Paths*) em redes IP entre pares de nós da rede cuja função é semelhante aos VPCs nas redes ATM. Assim, é importante investigar a possibilidade de usar os modelos propostos para o dimensionamento de redes IP com MPLS.

ANEXO A

RESOLUÇÃO DE EQUAÇÕES (TRÁFEGO PONTO-MULTIPONTO)

Seja q um valor real positivo menor do que 1. Seja m um valor inteiro positivo, b um valor inteiro positivo e g um valor real positivo tal que:

$$0 < g < b \leq m$$

Seja e uma variável inteira e p uma variável real positiva. Considerando as seguintes funções:

$$Cauda(b,m,e,p) = \frac{\sum_{n=b+1}^{m}\left(\frac{C_n^m}{n^{e-1}\cdot p^{n-1}}\right)}{\sum_{n=1}^{m}\left(\frac{C_n^m}{n^{e-1}\cdot p^{n-1}}\right)} \quad \text{(A.1)}$$

$$Med(b,m,e,p) = \frac{\sum_{n=1}^{b}\left(n\cdot\frac{C_n^m}{n^{e-1}\cdot p^{n-1}}\right)}{\sum_{n=1}^{b}\left(\frac{C_n^m}{n^{e-1}\cdot p^{n-1}}\right)} \quad \text{(A.2)}$$

pretende-se calcular as variáveis e e p tal que:

$$Cauda(b,m,e,p) \approx q \quad \text{(A.3)}$$

$$Med(b,m,e,p) = g \quad \text{(A.4)}$$

O método usado para resolver estas duas equações é dado por um processo iterativo em que se resolve primeiro a equação (A.4) e depois se verifica se a equação (A.3) é cumprida. Na primeira iteração, fixa-se um valor de e e calcula-se o parâmetro p dado pela equação (A.4). Com os dois parâmetros e e p, calcula-se $Cauda(b,m,e,p)$ e incrementa-se ou decrementa-se o valor de e consoante o valor de $Cauda(b,m,e,p)$ está abaixo ou acima de q. Este processo é repetido até ser obtido um par de valores e e p tal que e é máximo e o valor de $Cauda(b,m,e,p)$ seja menor ou igual que q. O método implementado pode ser descrito da seguinte forma:

Passo 1: Fazer $i = 1$ e $e_i = 0$

Passo 2: Calcular p_i tal que $Med(b,m,e_i,p_i) = g$

Passo 3: Calcular $q_i = Cauda(b,m,e_i,p_i)$

Se primeira iteração, então $q_{i-1} = q_i$

Passo 4: Se ($q_i \leq q$ e $q_{i-1} \leq q$) , então fazer $e_{i+1} = e_i + 1$ e $i = i + 1$ e saltar para Passo 2

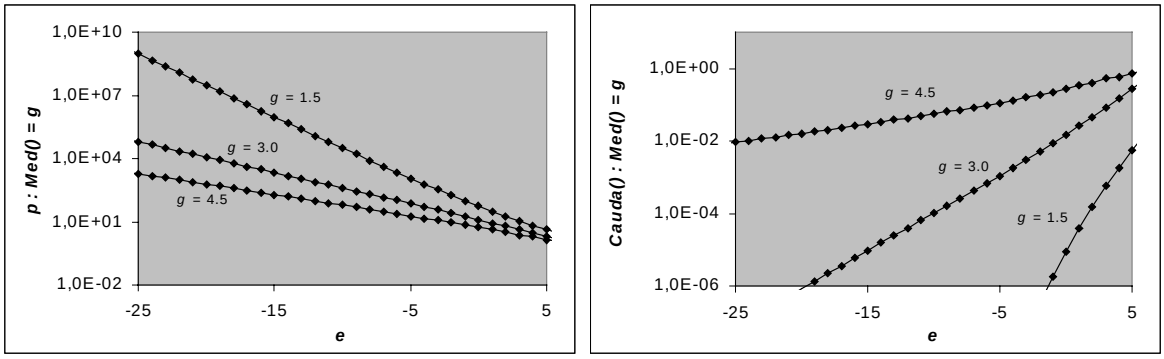
Se ($q_i \geq q$ e $q_{i-1} \geq q$) , então fazer $e_{i+1} = e_i - 1$ e $i = i + 1$ e saltar para Passo 2

Senão então: Se ($q_i < q_{i-1}$) então $e = e_i$ e $p = p_i$

Se ($q_{i-1} < q_i$) então $e = e_{i-1}$ e $p = p_{i-1}$

O passo 1 e o passo 3 são triviais. Consideremos primeiro o passo 4 e deixemos para o fim a implementação do passo 2. O passo 4 faz o seguinte. Se os dois últimos valores de e (que são dois inteiros contíguos) resultam ambos num valor de $Cauda(b,m,e,p)$ menor ou igual do que o seu valor objectivo q (equação A.3), então continuamos a incrementar o valor de e. Se pelo contrário, os dois últimos valores de e resultam ambos num valor de $Cauda(b,m,e,p)$ maior do que o seu valor objectivo q, então continuamos a decrementar o valor de e. No caso em que um valor de e faz o valor de $Cauda(b,m,e,p)$ ser maior ou igual que q e o outro faz o valor de $Cauda(b,m,e,p)$ ser menor ou igual que q, então escolhemos este último valor de e como solução final. Note-se que ao atribuirmos o valor inicial de zero a e no passo 1, não sabemos se, com o valor de p resultante do passo 2, vamos encontrar uma solução para $Cauda(b,m,e,p)$ inicialmente menor ou maior do que q. Por esta razão é que no passo 4 é preciso verificar se o valor de e precisa de ser decrementado ou incrementado até atingir o maior valor em que a função $Cauda(b,m,e,p)$ não ultrapasse o seu valor objectivo q. Este método é válido desde que a função $Cauda(b,m,e,p)$ seja

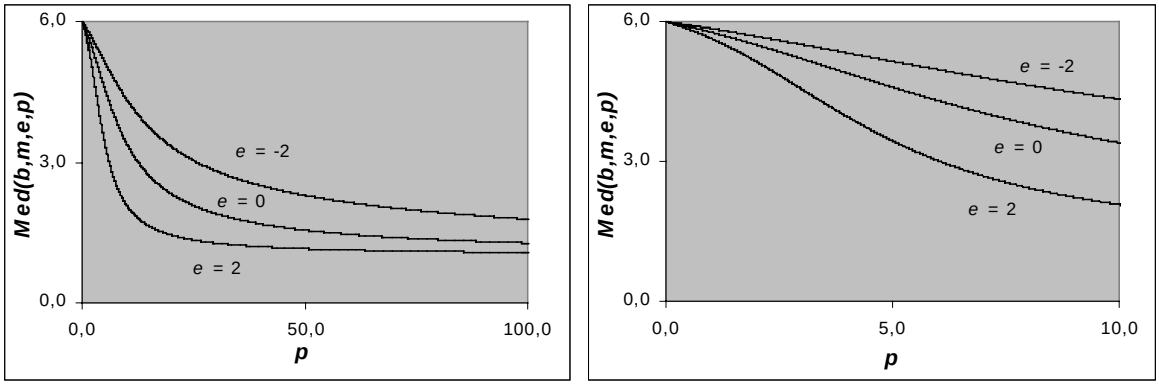
monotonicamente crescente em relação a *e* para valores de *e* e *p* tais que a função *Med(b,m,e,p)* seja igual a *g*.



(a) **(b)**

Figura A-1 : Para *m* = 29 e *b* = 6 (a) valor de *p* tal que *Med(b,m,e,p)* = *g*
(b) valor de *Cauda(b,m,e,p)* para valores de *e* e *p* tal que *Med(b,m,e,p)* = *g*

A Figura A-1 ilustra esta característica para valores de *m* = 29 e *b* = 6 e para 3 valores diferentes de *g*. A Figura A-1a mostra a variação do parâmetro *p* com o valor de *e* tal que a função $Med(b,m,e_i,p_i)$ é igual a *g*. Para cada um dos pares de valores (*e*,*p*) apresentados na Figura A-1a, a Figura A-1b mostra a função *Cauda(b,m,e,p)* resultante em que se verifica que é monotonicamente crescente com o parâmetro *e*. Não foi possível provar esta característica para todos os valores possíveis de *m*, *b*, *g* e *q* mas verificou-se que ela é válida para todos os casos de interesse estudados.



(a) **(b)**

Figura A-2 : *Med(b,m,e,p)* para *m* = 29 e *b* = 6 em função do parâmetro *p*
entre 0 e 100 (a) e entre 0 e 10 (b)

Abordemos agora a forma de resolver o passo 2 do método apresentado, ou seja, o cálculo da equação (A.4) em que o valor da variável *e* é dado e pretende-se calcular a

variável *p*. A Figura A-2 ilustra a função *Med(b,m,e,p)* para o caso de $m = 29$ e $b = 6$. Como se pode verificar, esta função é monotónica decrescente e a sua segunda derivada é positiva para quase todos os valores de *p* excepto nos casos em que o valor da função se aproxima de *b* (o limite de *Med(b,m,e,p)* quando *p* tende para zero é *b*) como ilustrado na Figura A-2b. Nos casos de interesse, *g* é substancialmente menor do que *b* pelo que estamos na região em que a segunda derivada da função *Med(b,m,e,p)* é positiva. Assim, a equação A.4 pode ser resolvida usando o método de Newton. Partindo de um valor inicial p_0, calcula-se iterativamente p_i segundo a seguinte fórmula em que $Med(p)'$ representa a derivada da função *Med()* no ponto *p*:

$$p_{i+1} = p_i + \frac{g - Med(p_i)}{Med(p_i)'}$$

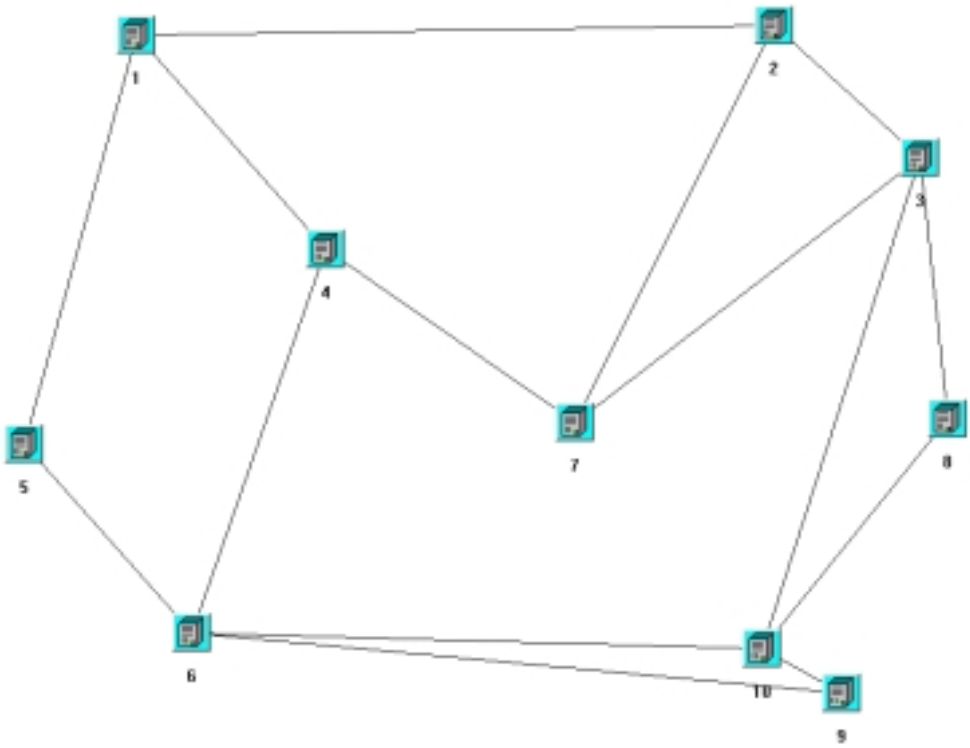
e termina-se o processo quando o valor de $g - Med(p_i)$ for suficientemente próximo de zero (no nosso caso menor que 1.0E−4). O método de Newton funciona com a condição de p_0 ser tal que $Med(p_0) > g$. Para o garantir, inicia-se p_0 com um valor arbitrariamente alto (no nosso caso 1000) e vai-se dividindo o valor por 2 até que a condição seja válida.

ANEXO B

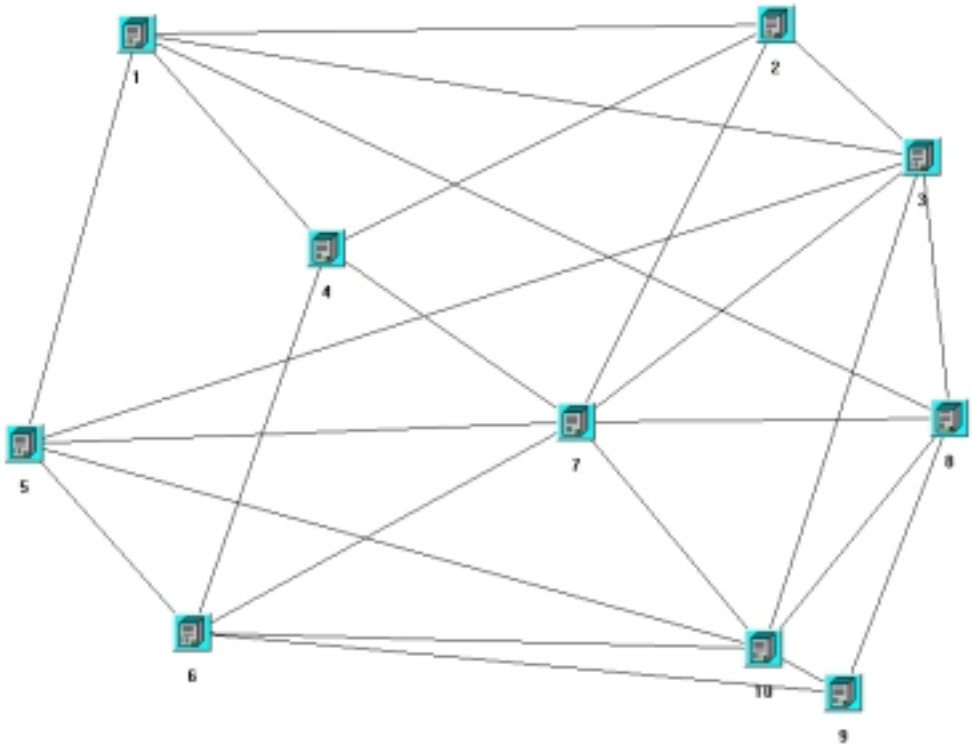
GRAFOS E RESULTADOS DO CAPÍTULO 3

Neste anexo são apresentados em primeiro lugar os grafos das redes que serviram de suporte aos resultados computacionais apresentados no capítulo 3 e, de seguida, os resultados computacionais de todos os problemas calculados para todos os algoritmos estudados. Para cada um dos problemas calculados, é apresentado o custo da melhor solução encontrada pelo respectivo algoritmo, o limite inferior do custo óptimo e o tempo de execução do algoritmo. Todos os problemas foram resolvidos numa plataforma PC com processador Pentium II a 350 MHz e com 256 Mbytes de memória RAM. O tempo de execução não inclui o cálculo preliminar das capacidades dos VPCs nem, para os modelos com trajectos candidatos, a determinação dos trajectos candidatos dos VPCs. Este tempo é, no entanto, desprezável em relação à resolução do modelo de optimização (nunca superior a 2%). Nas tabelas com os resultados computacionais, D.R. significa “designação da rede” e N.T.C. significa “número de trajectos candidatos”.

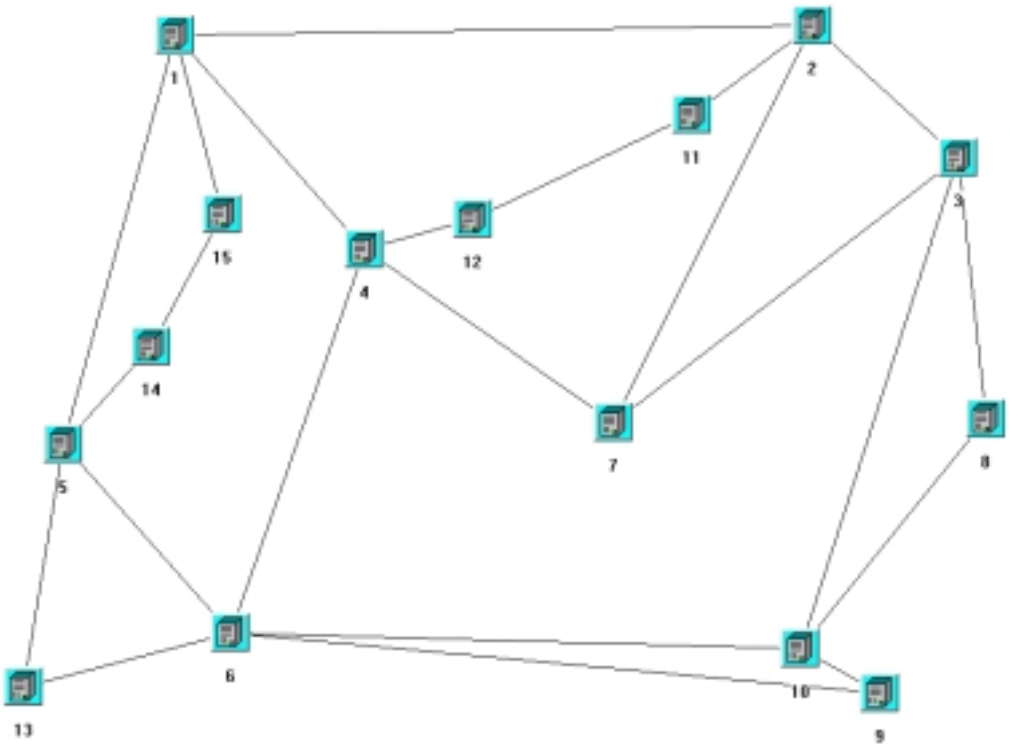
Rede 10S



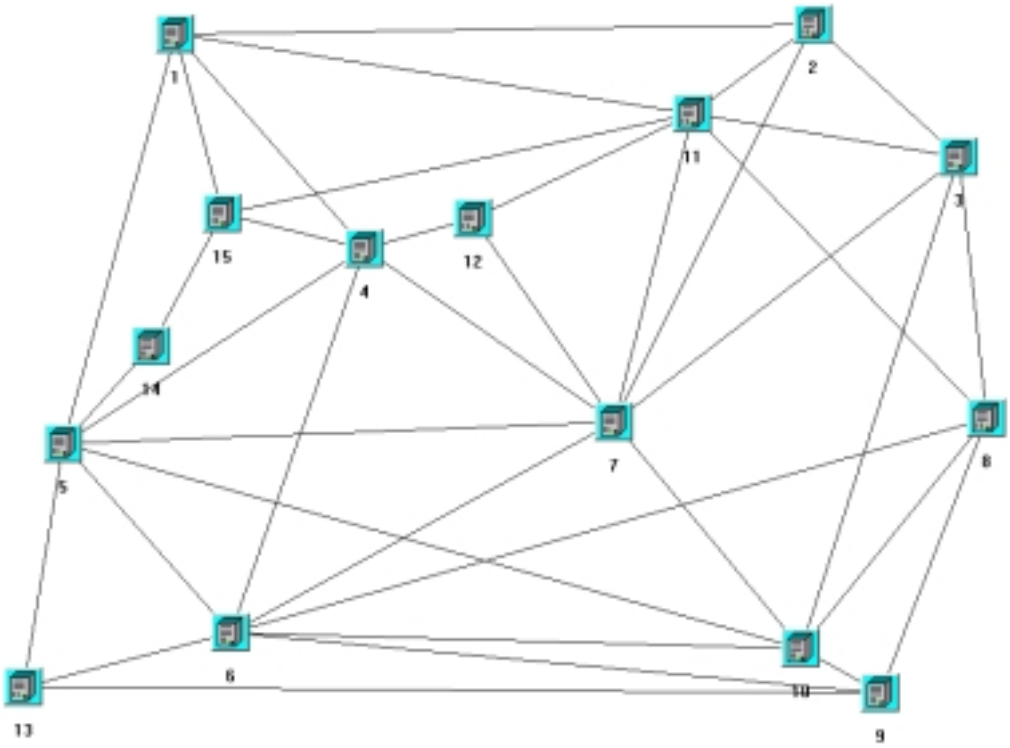
Rede 10D



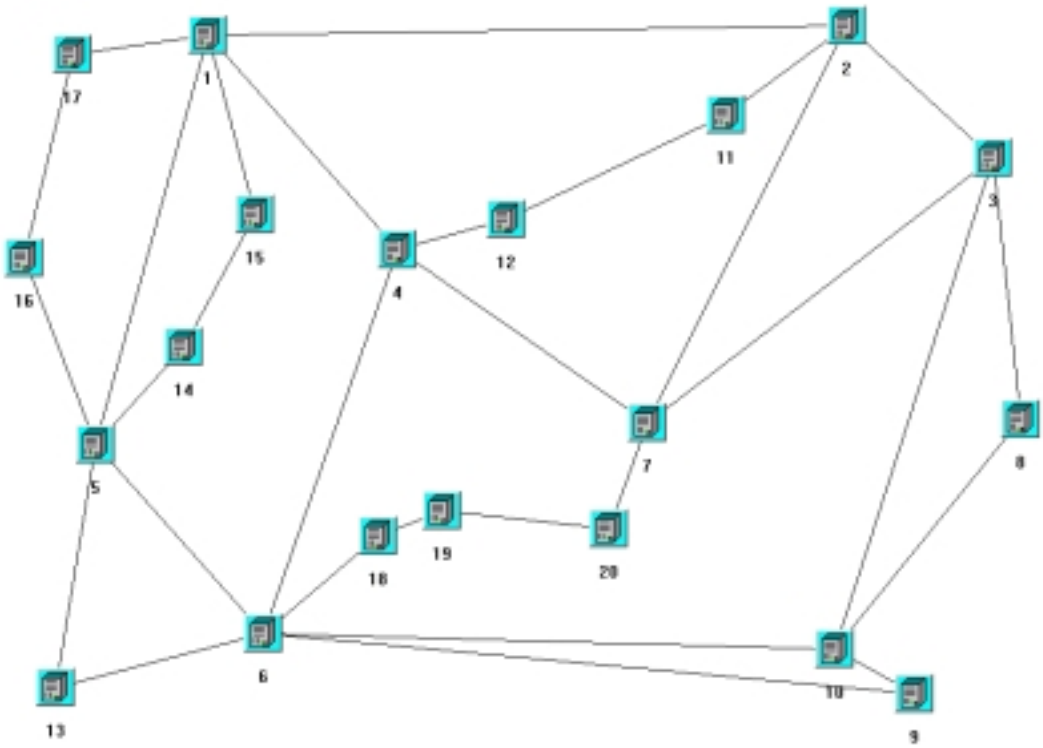
Rede 15S



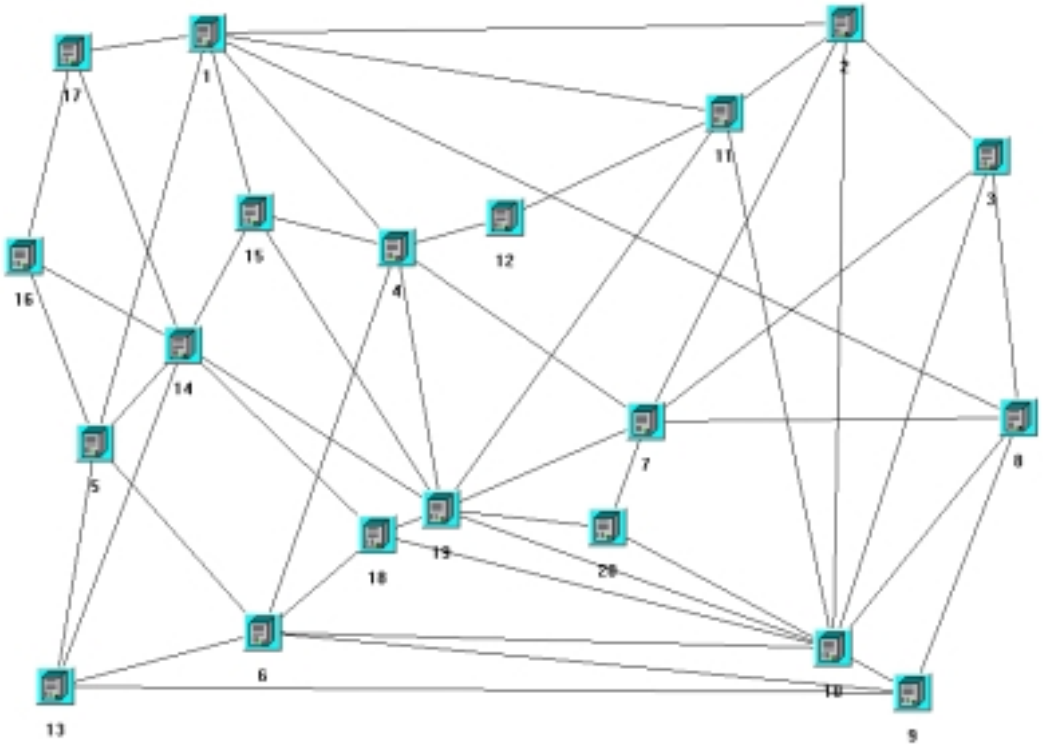
Rede 15D



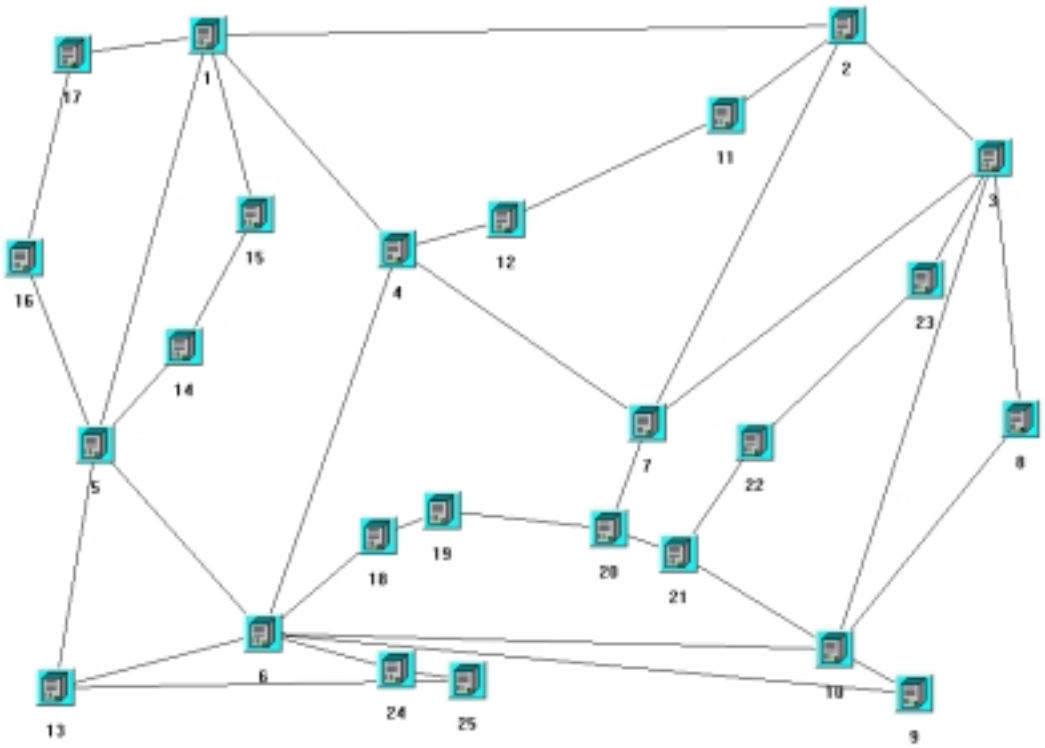
Rede 20S



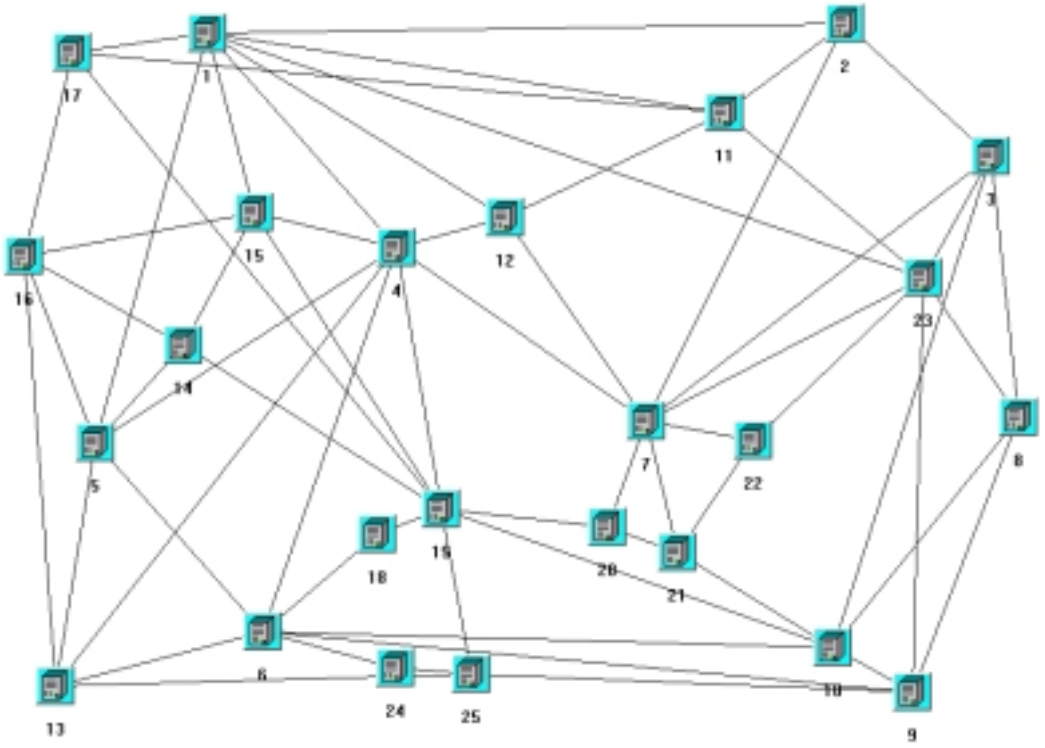
Rede 20D



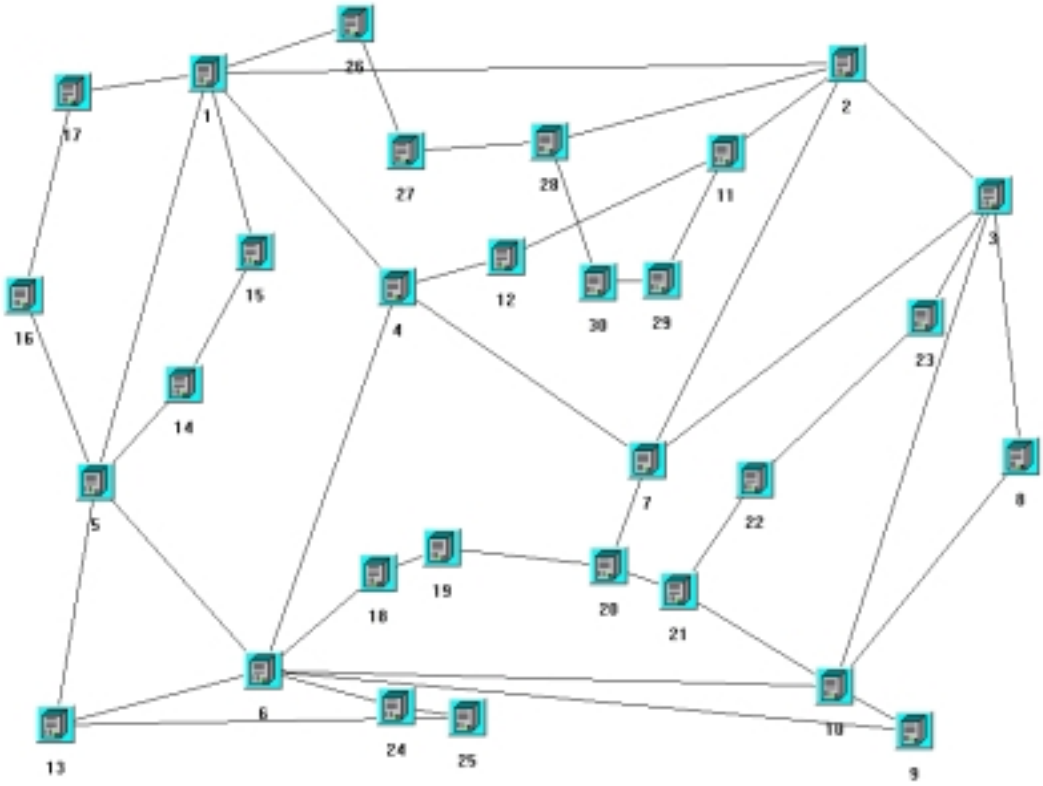
Rede 25S



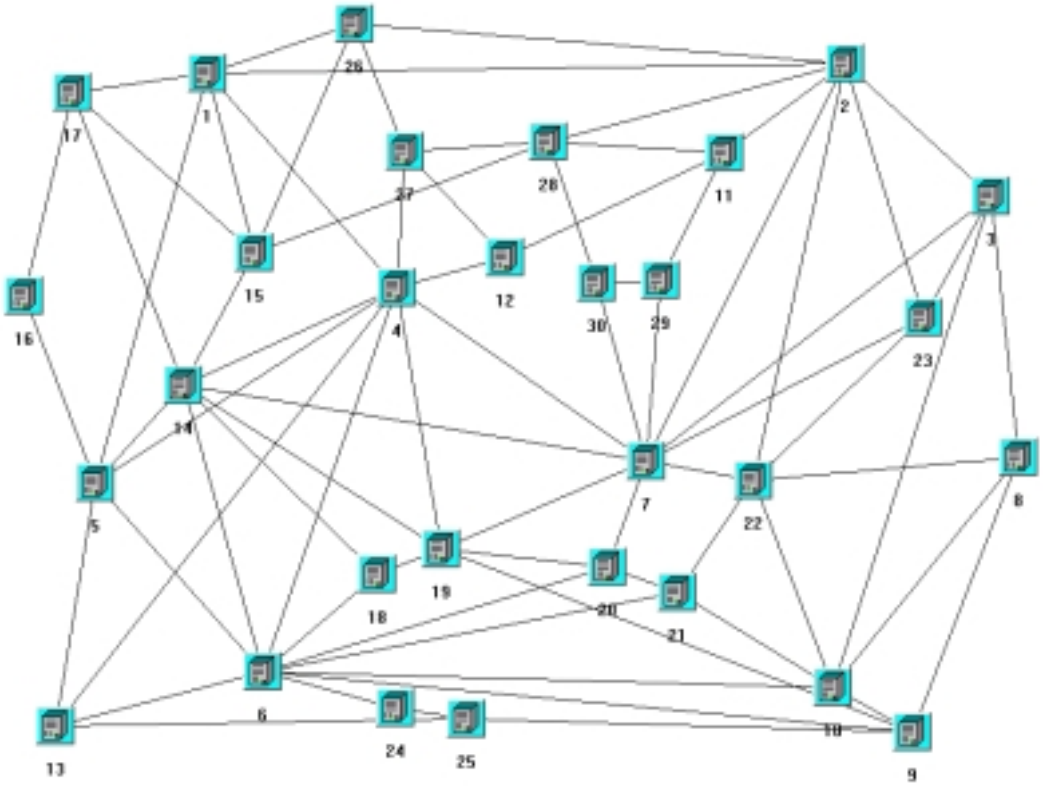
Rede 25D



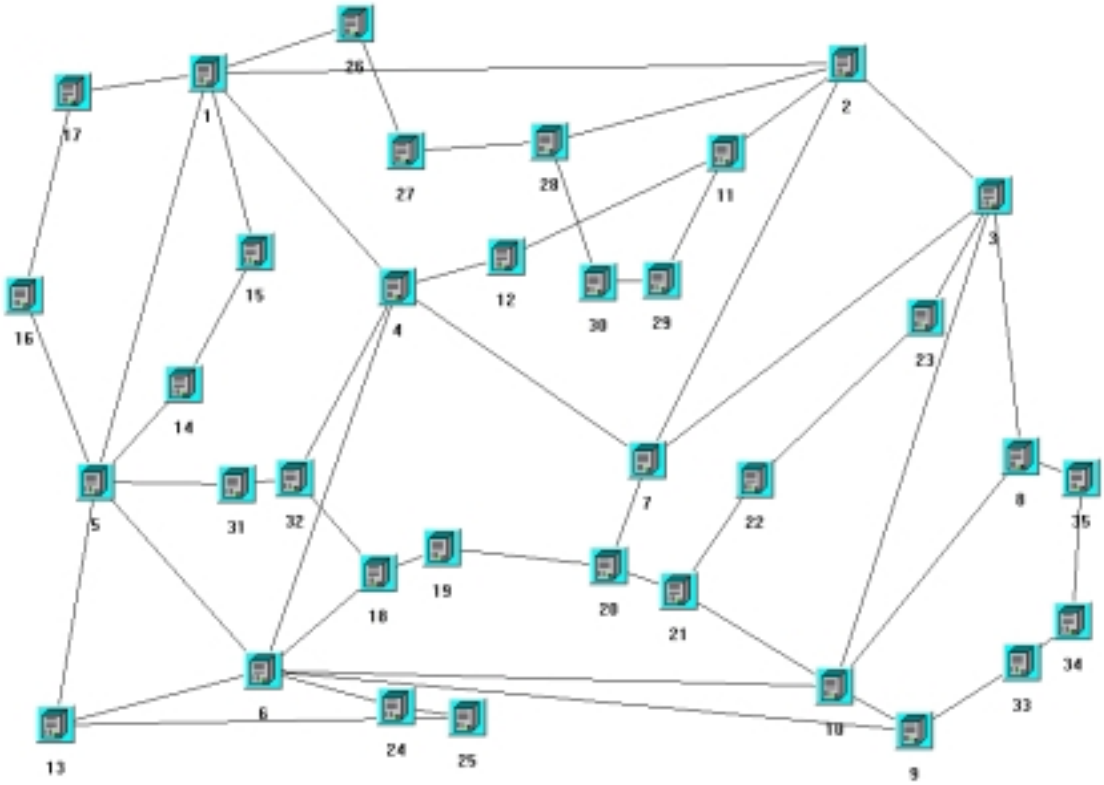
Rede 30S



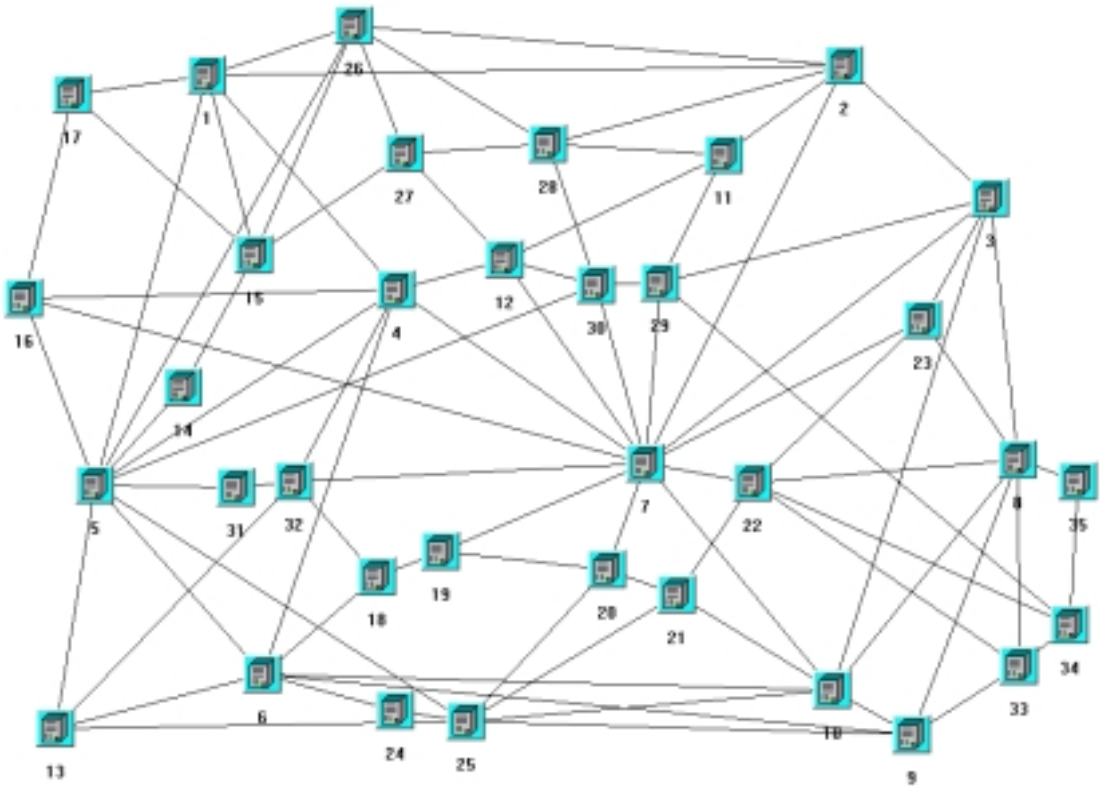
Rede 30D



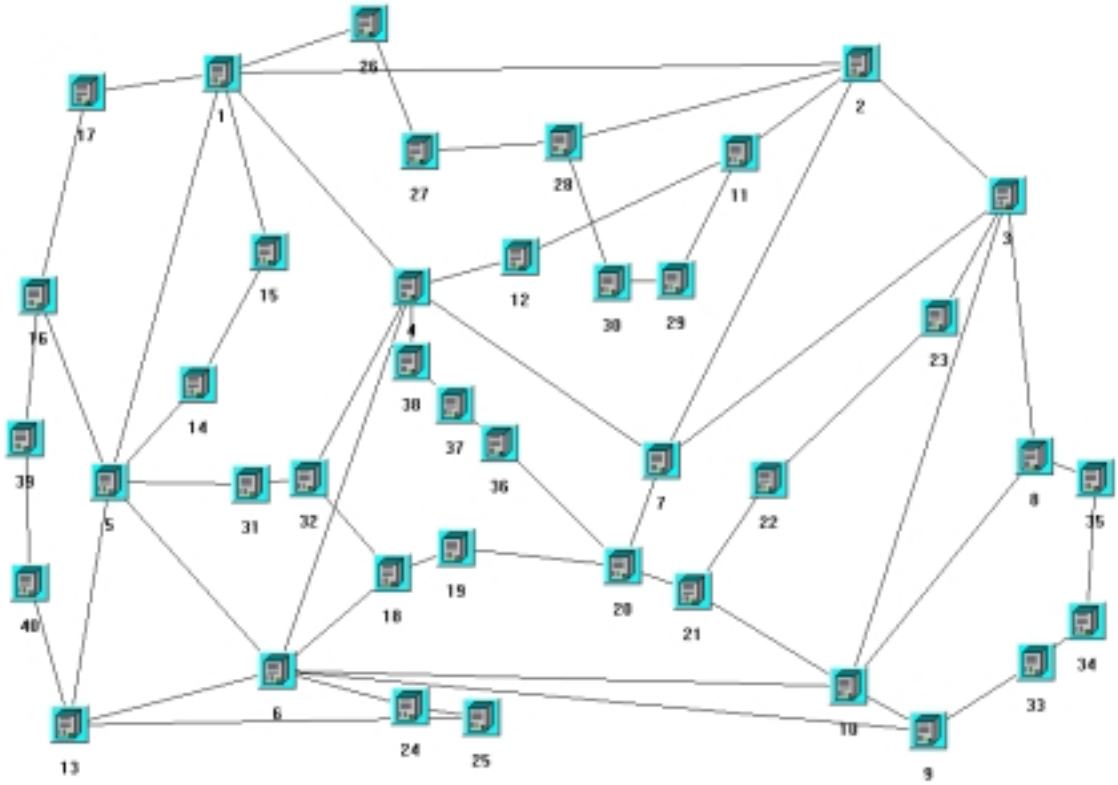
Rede 35S



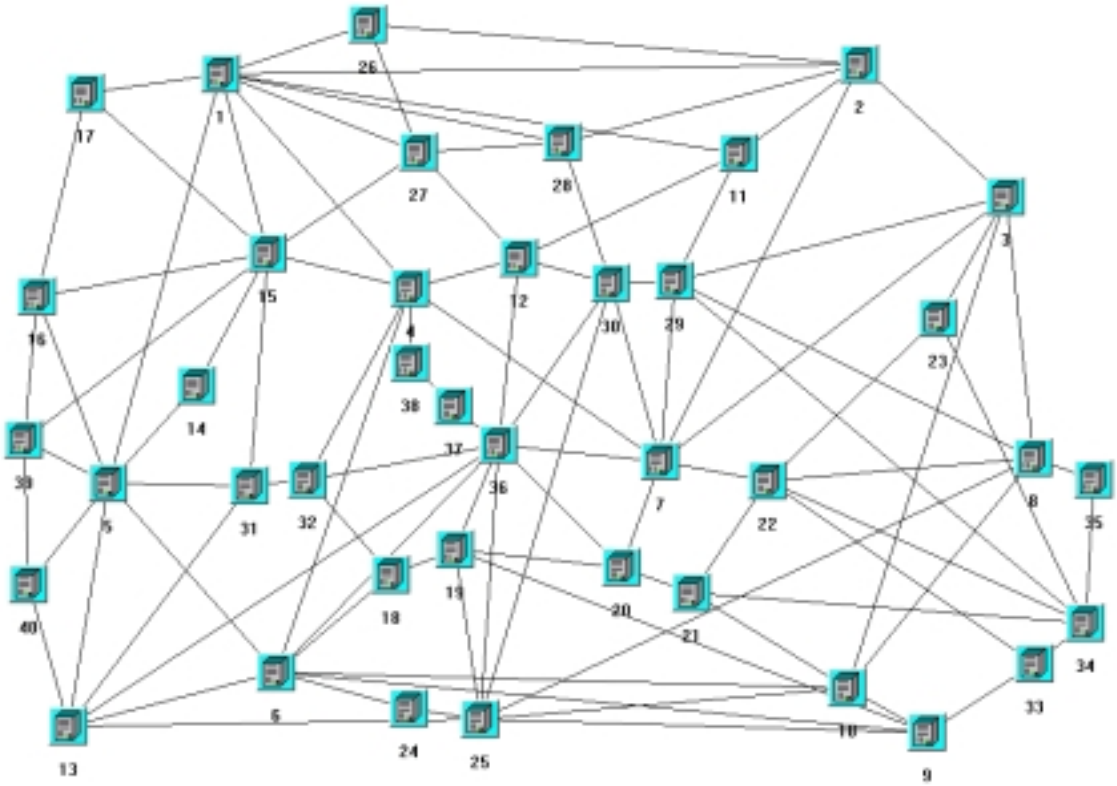
Rede 35D



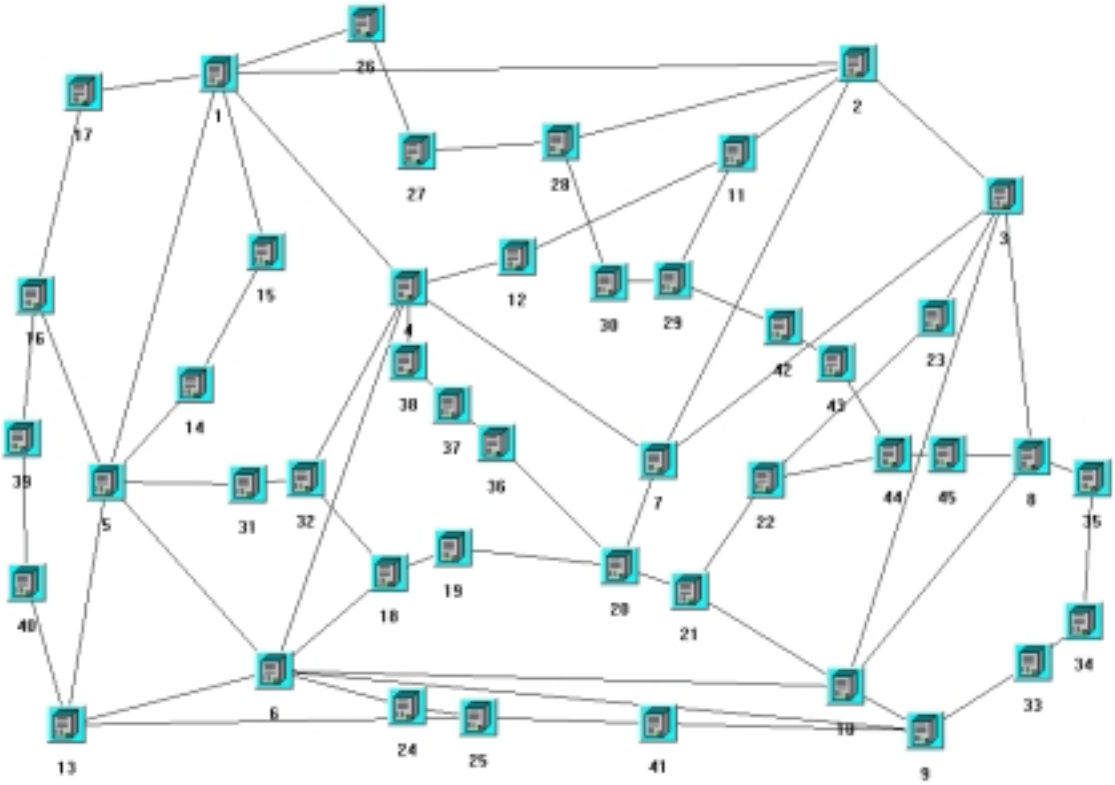
Rede 40S



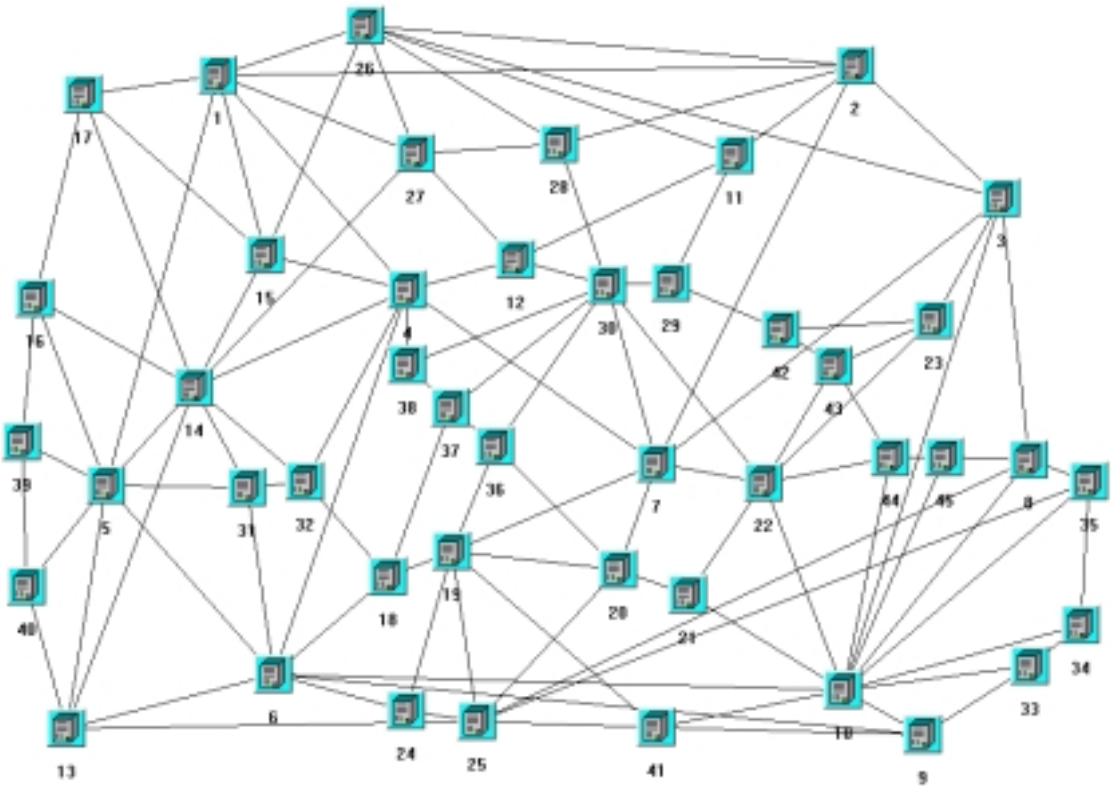
Rede 40D



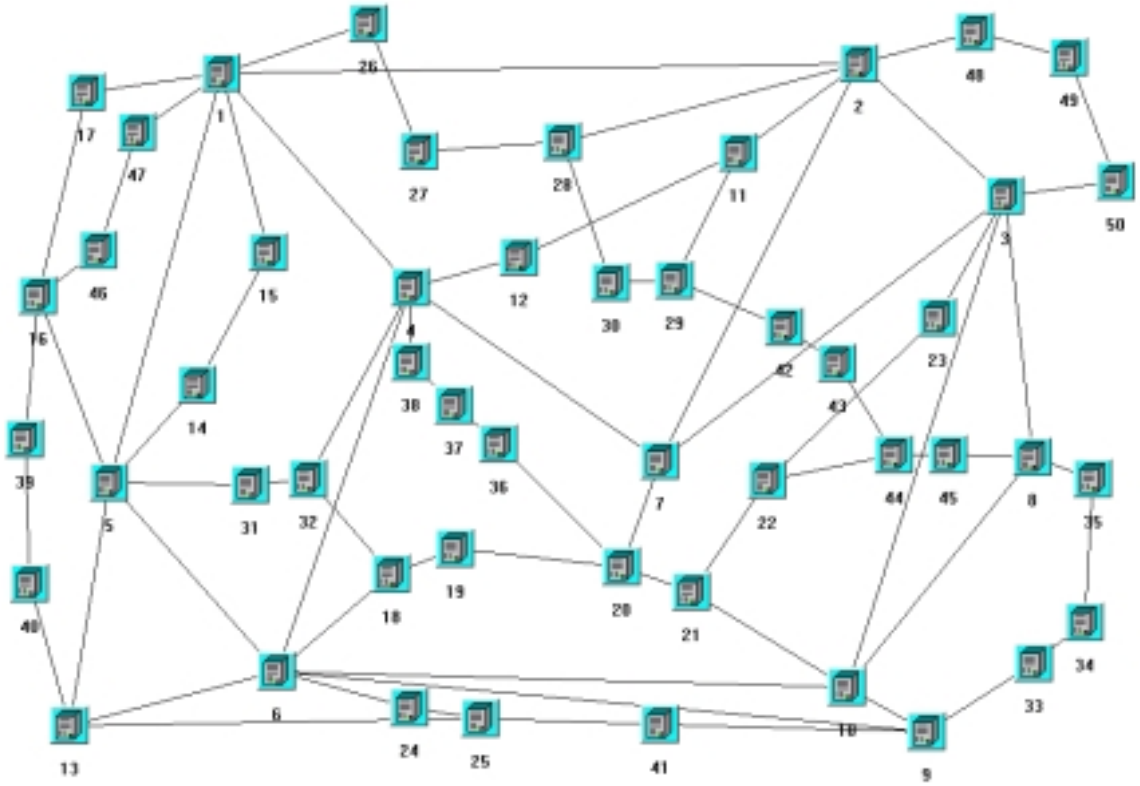
Rede 45S



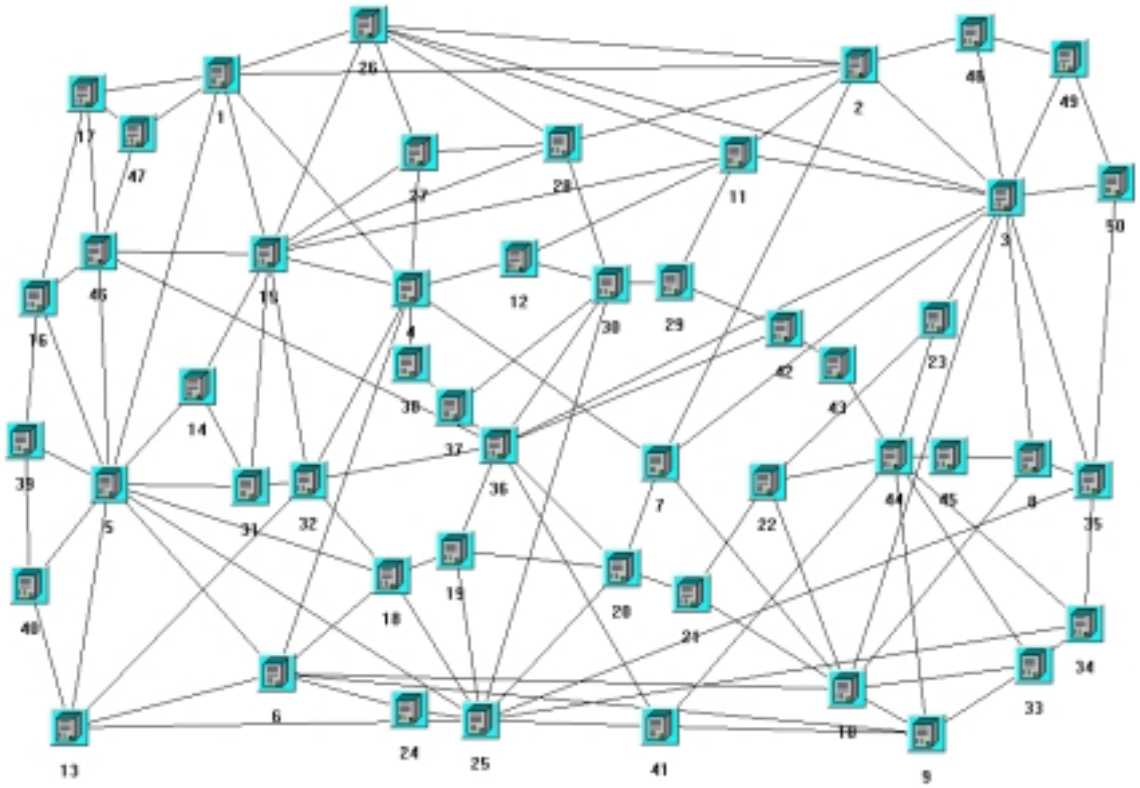
Rede 45D



Rede 50S



Rede 50D



Modelo por Formulação Explícita de VPC Uni-Horário

Custo da Melhor Solução

D.R.	1	2	3	4	5	6	7	8	9	10
10S	3012,6	2967,7	2900,0	2966,4	3030,3	2942,2	3186,0	2851,1	3042,0	3164,0
15S	6248,1	6367,4	6046,2	6154,2	6023,0	5688,9	6043,6	5526,8	5919,1	6109,2
20S	10249,9	10589,8	10524,1	10063,5	9494,5	10050,6	10017,9	10111,4	10297,0	9913,8
25S	16522,0	15858,8	16114,8	16155,9	15998,9	16226,7	16435,8	16198,1	15784,5	15928,6
30S	23008,6	24336,1	24078,3	24340,7	24133,0	23929,0	23558,4	23334,2	24315,1	23474,2
35S	33357,8	32965,1	34061,6	32793,3	33056,9	32907,4	33248,6	33781,0	32928,8	34269,0
40S	44635,7	43007,7	42808,3	43719,2	43255,2	43334,1	42457,0	44232,5	44082,1	44184,7
45S	59216,3	59020,1	59500,9	59232,1	58031,7	57699,8	58074,1	58826,1	58194,2	59834,0
50S	72207,8	72478,3	72796,1	72735,9	73082,3	73716,5	71726,7	72318,9	71827,0	73048,8

D.R.	1	2	3	4	5	6	7	8	9	10
10D	2806,1	2785,6	2832,4	2856,4	2905,2	2865,6	2848,9	2774,9	2913,3	2921,2
15D	6317,3	5905,7	6373,2	5909,5	5462,7	5894,5	6054,6	6195,7	5873,2	6202,6
20D	11295,0	11333,7	11190,8	11549,5	10414,2	11297,2	11819,3	10793,2	11238,4	10881,0
25D	16392,6	17206,4	16210,4	16735,9	16589,9	16756,4	16263,2	15826,2	16458,8	16518,3
30D	22752,4	21656,6	22431,5	22498,9	22382,8	22411,1	22107,1	22624,3	22292,0	22456,8
35D	29844,6	30140,9	30746,4	30119,9	30150,3	29875,9	30009,6	30166,4	29458,6	30486,4
40D	38971,9	39828,6	40139,8	38221,1	39444,4	38357,6	38835,8	39750,1	39740,7	40113,6
45D	49815,1	49819,3	50218,8	49819,0	49488,1	51391,5	49221,6	50294,6	51152,8	50385,0
50D	59142,6	60030,1	59635,0	59614,0	59664,1	59543,9	59670,5	59208,3	60448,9	59302,8

Limite Inferior do Custo da Solução Óptima

D.R.	1	2	3	4	5	6	7	8	9	10
10S	1454,7	1450,9	1400,9	1441,5	1498,6	1366,4	1509,6	1349,9	1437,3	1429,1
15S	3991,4	3934,8	3778,0	3825,3	3943,6	3616,3	3825,9	3472,0	3690,4	3766,8
20S	7802,6	7922,0	7874,6	7473,2	7254,2	7803,1	7516,1	7667,6	7561,8	7474,4
25S	13429,1	12739,4	12910,2	13305,2	12956,9	13178,3	13033,7	12933,2	12818,8	12861,9
30S	19638,1	20603,5	20894,6	20884,3	20365,4	20366,7	20326,4	19870,3	20869,9	19889,9
35S	29333,6	29243,5	29667,5	28784,1	29471,5	28915,0	29234,9	30032,6	29098,7	30023,0
40S	39853,7	38408,9	38301,9	39190,7	39069,4	38871,1	38166,6	39801,3	39613,0	39782,9
45S	53694,6	53579,6	54575,5	54057,4	53028,0	52933,9	52926,0	53668,1	52864,0	54794,5
50S	67074,1	67107,9	67581,2	67596,5	67819,0	68509,5	66705,6	66671,6	66812,2	67736,0

D.R.	1	2	3	4	5	6	7	8	9	10
10D	1169,2	1063,1	1136,6	1034,5	1127,0	1201,4	1196,6	1064,3	1251,0	1098,2
15D	2970,2	2978,2	3002,4	2894,1	2915,0	2990,4	3060,8	3144,4	3028,6	3021,5
20D	5866,0	6118,2	5762,6	6039,8	6048,6	6105,9	6355,2	5890,5	5759,7	5595,7
25D	10135,6	10652,9	10148,0	10198,3	10073,7	10206,1	9714,5	9658,8	10439,2	10113,8
30D	16049,4	15158,3	15857,9	15376,6	15900,5	15710,5	15356,1	15765,9	15487,0	15805,6
35D	22112,9	22403,0	22868,2	22234,9	22308,0	22680,8	21969,6	22152,9	22032,7	22615,5
40D	30798,3	30931,8	31981,0	30030,3	31213,6	30011,8	31040,5	31688,9	31480,2	31689,4
45D	41420,4	40413,8	40856,8	40319,1	40605,3	42057,1	40449,0	41113,5	41791,1	41007,7
50D	49532,6	50113,7	49396,7	49636,5	49487,2	49178,2	49425,5	49508,8	50305,0	49349,5

Tempo de Cálculo (segundos)

D.R.	1	2	3	4	5	6	7	8	9	10
10S	1,12	1,11	1,15	1,11	1,10	1,10	1,11	1,12	1,12	1,10
15S	3,12	3,10	3,15	3,03	3,03	3,07	3,03	3,07	3,11	3,01
20S	6,76	6,80	6,81	6,78	6,68	6,80	6,79	6,78	6,78	6,78
25S	12,20	12,54	12,26	12,34	12,30	12,28	12,35	12,26	12,34	12,27
30S	20,12	20,17	20,19	20,34	20,19	20,33	20,28	20,09	20,36	20,17
35S	31,31	31,22	31,23	31,21	31,30	31,11	31,08	31,56	31,24	31,13
40S	46,81	45,70	45,99	46,39	46,72	46,23	46,10	46,69	46,41	46,46
45S	67,70	68,67	65,54	68,97	65,73	65,58	66,14	66,81	68,19	66,33
50S	90,30	90,55	88,64	89,84	89,51	90,20	90,35	89,02	90,24	89,73

D.R.	1	2	3	4	5	6	7	8	9	10
10D	1,28	1,31	1,27	1,30	1,31	1,29	1,46	1,38	1,30	1,29
15D	3,76	3,82	3,75	3,79	3,80	3,83	3,80	3,77	3,78	3,79
20D	8,54	8,57	8,55	8,57	8,65	8,49	8,59	8,61	8,61	8,56
25D	15,93	16,05	15,85	15,89	15,95	15,83	15,88	15,99	15,96	15,85
30D	26,30	26,24	26,17	26,43	26,36	26,22	26,16	26,42	26,24	26,31
35D	40,94	40,95	41,25	41,08	41,26	41,54	41,11	41,22	41,08	41,22
40D	63,27	63,68	62,15	63,50	63,82	61,45	62,59	63,11	63,40	64,08
45D	91,04	91,17	90,90	91,61	91,21	92,43	91,36	91,84	91,65	92,29
50D	122,58	122,62	120,49	120,38	119,93	120,03	120,39	118,99	121,43	120,95

Modelo por Formulação Explícita de VPC Multi-Horário

Custo da Melhor Solução

D.R.	1	2	3	4	5	6	7	8	9	10
10S	2754,4	2835,2	2698,1	2856,8	2875,7	2480,2	2759,3	2561,9	2611,8	2694,0
15S	5409,7	5321,0	5200,5	5642,1	5141,7	5040,5	5421,7	5137,7	5345,4	5336,7
20S	8855,6	8762,9	8854,5	8495,9	8474,1	8703,3	8414,9	8595,8	8405,6	8463,5
25S	13291,7	12593,0	12928,9	12850,2	12369,9	13087,9	13174,9	12887,4	12533,6	12456,8
30S	17959,4	18397,4	18640,4	18327,0	18864,6	18380,2	18681,3	18093,6	18819,2	18281,1
35S	25157,8	24867,5	26096,5	25033,6	25049,4	24665,0	24873,8	25860,0	24938,4	26001,9
40S	33712,6	31497,3	32174,7	32803,3	32219,9	32670,3	32209,7	32813,1	33353,0	33372,3
45S	43061,4	43774,5	44376,2	43982,7	42049,3	42154,4	43077,1	43760,1	42772,6	44084,6
50S	53678,3	53562,7	54165,3	53853,9	54277,3	54079,0	53092,6	52539,1	53052,5	53467,2

D.R.	1	2	3	4	5	6	7	8	9	10
10D	2367,0	2484,0	2755,4	2599,2	2831,0	2653,3	2789,0	2426,8	2778,7	2529,4
15D	5110,2	5315,5	5128,5	5271,1	5005,4	5124,4	5417,9	5224,6	5246,9	4958,1
20D	8344,7	8832,7	8592,7	8845,0	9331,7	8815,8	9282,6	8575,2	8790,8	8182,9
25D	14012,8	13655,4	13904,4	13775,1	14604,8	14105,7	13001,7	13400,7	14210,9	14115,0
30D	19648,0	19154,3	19167,8	19248,7	19674,7	19421,1	18904,8	19337,5	19383,7	19411,6
35D	24549,6	25216,3	25110,1	24524,4	25135,6	24570,1	24700,5	24554,8	24564,6	24989,5
40D	31990,3	31297,5	32271,2	31046,1	32199,1	31387,2	32444,7	32253,4	32401,0	32324,3
45D	39095,8	38537,9	40040,6	38980,3	40668,8	40203,5	38479,3	39771,4	39956,8	39560,2
50D	46665,1	47442,3	46242,9	46038,9	46192,0	46738,2	46217,3	45883,0	47052,0	46912,1

Limite Inferior do Custo da Solução Óptima

D.R.	1	2	3	4	5	6	7	8	9	10
10S	968,8	859,9	925,1	973,2	938,7	940,9	928,6	918,1	820,7	916,9
15S	2744,0	2578,3	2452,8	2548,7	2490,0	2461,8	2537,4	2292,0	2478,5	2465,1
20S	5371,2	5280,1	5519,7	5328,3	5075,3	5299,2	5188,6	5281,8	5179,6	5207,5
25S	9337,3	8852,9	9000,2	9253,1	8812,8	9215,6	9055,0	9061,7	8932,9	8943,9
30S	13694,7	14239,2	14388,8	14386,2	14328,4	14219,3	14258,9	13672,8	14640,0	13688,3
35S	20368,3	20178,2	21065,5	19865,9	20283,5	19779,9	19986,1	20993,8	19961,5	20980,3
40S	27792,8	26592,7	26719,2	26962,0	26844,4	26908,3	26319,3	27071,3	27236,2	27200,1
45S	36976,5	37200,2	37648,3	37648,9	36724,0	35992,5	36383,6	37287,8	36663,5	37664,2
50S	46404,6	46640,1	46695,5	46606,2	47005,9	46882,0	45979,7	45467,0	46166,7	46376,3

D.R.	1	2	3	4	5	6	7	8	9	10
10D	703,9	740,9	756,6	645,1	676,5	782,2	816,8	714,2	792,0	710,2
15D	1909,2	2067,4	1731,2	1986,3	1877,5	2000,1	2111,6	2137,3	2085,0	1902,8
20D	4116,6	4045,3	3956,8	4238,4	4126,1	4185,0	4537,0	4113,2	3530,8	3940,8
25D	6816,7	7317,1	7083,6	6922,4	6970,5	6959,8	6578,5	6676,5	7063,1	6950,9
30D	11209,7	10447,3	10895,2	10736,9	11053,0	10790,8	10461,9	10806,1	10744,1	10958,2
35D	15066,7	15544,2	15544,4	15356,3	15294,7	15593,2	15055,7	15224,0	15314,2	15545,5
40D	21100,7	21072,2	21849,0	20629,0	21314,7	20872,4	21480,4	21620,6	21554,4	21757,3
45D	28343,1	27726,0	27672,1	27650,9	27988,0	29129,0	27625,2	28142,9	28933,7	28220,1
50D	34039,3	34218,3	33831,0	33850,0	33651,5	34052,2	33746,4	33643,5	34420,6	33701,4

Tempo de Cálculo (segundos)

D.R.	1	2	3	4	5	6	7	8	9	10
10S	3,29	3,29	3,31	3,29	3,30	3,30	3,29	3,32	3,31	3,37
15S	9,41	9,33	9,34	9,24	9,43	9,36	9,22	9,19	9,19	9,47
20S	20,01	20,24	20,23	20,12	19,94	19,96	20,12	20,40	19,88	20,15
25S	37,20	38,16	37,99	37,70	37,55	37,87	37,74	37,54	37,74	37,99
30S	62,88	63,08	66,80	63,04	63,78	63,45	63,16	62,67	63,92	63,57
35S	98,17	97,87	99,07	98,77	98,24	97,91	98,69	99,39	98,28	98,38
40S	144,49	143,55	144,27	144,20	143,99	144,62	144,23	144,61	145,01	144,41
45S	203,80	204,40	205,72	208,40	206,79	205,68	207,18	205,52	206,30	208,01
50S	283,50	281,90	283,10	282,11	281,10	284,76	281,30	283,10	283,10	281,47

D.R.	1	2	3	4	5	6	7	8	9	10
10D	3,84	3,81	3,83	3,83	3,82	3,89	3,86	3,83	3,84	3,87
15D	11,68	11,79	11,72	11,86	11,87	11,81	11,56	11,81	11,79	11,77
20D	26,78	26,30	26,45	26,14	27,01	26,55	26,86	26,13	26,81	26,84
25D	49,90	49,63	49,47	49,47	49,44	49,34	49,42	49,54	49,81	50,27
30D	83,73	83,27	82,84	83,39	83,00	82,59	83,00	83,22	82,56	83,43
35D	131,61	131,26	131,70	131,45	133,96	132,10	131,04	131,16	131,49	131,08
40D	193,34	193,09	193,98	195,26	194,46	192,70	196,96	195,62	194,25	197,12
45D	283,10	278,70	292,64	288,50	290,45	279,60	275,85	288,30	278,07	288,61
50D	381,60	383,99	383,70	381,44	383,03	383,14	386,27	382,99	386,17	386,48

Modelo por Trajectos Candidatos de VPC Uni-Horário

Custo da Melhor Solução

D.R.	N.T.C.	1	2	3	4	5	6	7	8	9	10
10S	5	3300,4	3124,0	3083,9	3243,6	3351,4	3056,5	3194,3	3199,7	3003,9	3079,8
10S	20	2950,8	2923,6	2994,8	2938,9	3114,2	2724,1	3191,4	3160,3	2738,9	3024,7
15S	5	6440,9	6248,3	5516,6	6193,9	5934,6	5901,2	6213,1	5625,2	5961,7	6172,8
15S	20	6124,9	6345,5	6075,0	6262,3	6142,8	5971,8	6091,7	5386,4	6220,3	6221,9
30S	5	23015,0	24620,0	24317,3	24714,3	23898,3	23838,8	23521,6	23486,6	24229,8	23484,8
30S	20	23155,9	24451,4	24085,3	24686,4	23906,1	23827,8	23420,3	23388,6	24417,1	23373,7
45S	5	60150,0	60218,9	61413,5	60806,4	59857,3	59608,7	59785,8	60432,4	59648,6	61594,3
45S	20	58784,2	58669,4	59420,0	58943,1	58089,0	57531,1	58077,8	58808,4	58360,5	59694,7

D.R.	N.T.C.	1	2	3	4	5	6	7	8	9	10
10D	5	3461,3	3165,1	3478,4	3175,5	3438,9	3268,9	3108,2	3277,9	3610,6	3206,1
10D	20	3242,0	3085,0	2931,5	3224,0	3037,1	2885,1	2915,8	3071,1	3319,4	3251,7
15D	5	6958,7	6555,1	7160,9	6898,5	6551,6	6710,0	7036,4	6868,4	6579,1	7087,7
15D	20	6715,2	6079,9	6349,1	6271,6	5862,8	5911,2	6220,9	6359,8	5791,4	6300,9
30D	5	23773,0	22639,4	23212,9	23230,1	23619,1	23932,6	23289,9	23445,1	23453,2	23713,8
30D	20	22928,4	22298,0	22711,6	22602,0	22887,7	22788,8	22516,6	22614,4	22783,0	22781,6
45D	5	54979,4	54512,0	54339,9	54347,4	54026,6	55828,3	53620,5	55334,2	55425,3	54828,8
45D	20	52319,3	51412,0	51988,4	51500,9	51590,7	53721,7	51419,0	52901,6	53168,6	52056,5

Limite Inferior do Custo da Solução Óptima

D.R.	N.T.C.	1	2	3	4	5	6	7	8	9	10
10S	5	1454,8	1450,8	1400,8	1441,5	1498,5	1366,6	1509,7	1349,9	1437,3	1429,2
10S	20	1454,5	1451,0	1400,9	1441,5	1498,3	1366,4	1509,6	1349,9	1437,4	1429,0
15S	5	3991,3	3934,6	3777,5	3825,5	3943,8	3616,6	3825,6	3472,6	3690,7	3766,7
15S	20	3991,3	3934,5	3777,9	3825,5	3943,7	3617,1	3826,0	3472,7	3690,7	3766,7
30S	5	19716,9	20703,7	20972,8	20980,7	20471,3	20479,5	20405,9	19950,9	20925,9	19982,2
30S	20	19665,4	20632,4	20915,6	20914,6	20392,5	20399,6	20341,3	19885,9	20892,5	19911,8
45S	5	55381,6	55405,9	56395,7	55973,8	54429,4	54867,2	54783,2	55638,3	54714,3	56658,4
45S	20	53791,5	53713,1	54625,8	54191,3	53108,6	52976,9	52985,4	53756,4	53123,3	54868,5

D.R.	N.T.C.	1	2	3	4	5	6	7	8	9	10
10D	5	1169,2	1063,0	1136,6	1034,5	1127,0	1200,6	1196,7	1064,3	1250,9	1098,3
10D	20	1169,1	1063,0	1136,6	1034,5	1126,9	1201,4	1196,6	1064,3	1251,0	1098,3
15D	5	3000,9	2988,9	3051,8	2918,1	2919,3	3007,1	3074,4	3179,1	3052,6	3046,5
15D	20	2970,3	2977,8	3002,8	2894,2	2915,0	2990,5	3060,5	3144,7	3028,3	3021,6
30D	5	16710,2	15790,0	16514,3	16043,6	16638,3	16446,6	15939,7	16312,4	16161,2	16403,2
30D	20	16215,8	15338,2	16013,6	15572,8	16114,9	15933,1	15500,0	15897,3	15687,0	15949,6
45D	5	45902,1	45184,5	45068,0	45246,5	45010,1	46876,1	44280,0	45713,8	46131,1	45847,0
45D	20	43320,9	42631,1	42738,7	42461,5	42684,2	44323,3	42415,8	43313,0	43920,3	43308,2

Tempo de Cálculo (segundos)

D.R.	N.T.C.	1	2	3	4	5	6	7	8	9	10
10S	5	0,87	0,86	0,85	0,85	0,86	0,86	0,86	0,87	0,85	0,86
10S	20	2,41	2,39	2,39	2,39	2,44	2,41	2,38	2,38	2,41	2,48
15S	5	2,77	2,76	2,75	2,74	2,75	2,76	2,75	2,73	2,75	2,42
15S	20	8,60	8,41	8,36	8,68	8,65	8,72	8,74	8,57	8,69	8,74
30S	5	22,80	22,87	22,82	23,01	23,03	22,93	22,92	22,82	22,91	22,93
30S	20	73,20	73,15	72,94	73,34	73,26	73,08	73,20	73,31	73,13	73,39
45S	5	79,34	80,20	79,97	79,52	80,77	80,13	79,96	79,93	80,06	80,09
45S	20	242,10	240,30	240,51	241,01	241,38	241,35	241,32	240,99	241,51	241,20

D.R.	N.T.C.	1	2	3	4	5	6	7	8	9	10
10D	5	1,26	1,27	1,26	1,26	1,25	1,29	1,25	1,26	1,26	1,26
10D	20	3,72	3,73	3,73	3,73	3,72	3,79	3,76	3,75	3,74	3,80
15D	5	4,26	4,29	4,27	4,29	4,32	4,27	4,27	4,29	4,25	4,28
15D	20	13,37	13,19	13,19	13,65	13,83	13,65	13,57	13,69	13,52	13,62
30D	5	37,40	37,41	37,44	37,47	37,54	37,61	37,39	37,45	37,94	37,48
30D	20	113,30	114,24	114,50	114,37	114,34	114,31	114,29	114,32	114,19	114,22
45D	5	129,51	128,90	130,43	149,88	131,88	130,66	130,17	129,37	129,73	129,16
45D	20	385,42	387,97	387,80	388,42	389,14	389,80	389,22	388,81	389,16	388,44

Modelo por Trajectos Candidatos de VPC Multi-Horário

Custo da Melhor Solução

D.R.	N.T.C.	1	2	3	4	5	6	7	8	9	10
10S	5	2938,1	3020,2	2947,8	2938,1	3108,2	2814,6	3141,4	2791,3	3147,2	3120,2
10S	20	2938,1	3107,9	2947,8	2938,1	3108,2	2836,3	3175,8	2791,3	3035,0	3133,8
15S	5	5514,6	5973,5	5136,1	5893,8	5318,7	5355,3	5605,4	5667,2	5702,0	5169,1
15S	20	5163,3	5713,3	4768,3	6151,6	5157,5	5108,1	5472,0	4982,7	5100,4	5284,3
30S	5	18048,8	18632,2	18894,0	18522,7	19146,1	18846,6	18782,2	18181,8	19136,8	18124,7
30S	20	18083,8	18379,6	18654,5	18398,0	18897,1	18572,4	18695,1	18033,5	18982,9	18111,7
45S	5	44573,2	45318,8	45418,7	45363,7	43441,9	44077,7	44372,2	45599,1	44467,4	45961,2
45S	20	43242,5	44199,5	44091,2	44026,9	43469,0	41981,0	43121,4	43849,7	43006,9	43954,6

D.R.	N.T.C.	1	2	3	4	5	6	7	8	9	10
10D	5	3889,9	2945,2	3066,1	4211,2	3481,1	2793,4	3811,5	3346,5	3296,6	4205,6
10D	20	2689,5	2571,2	3121,7	3287,5	3065,9	2927,0	2986,1	3028,4	3478,2	4256,3
15D	5	6686,1	6260,3	7060,2	6262,2	6401,5	5512,1	6594,5	6986,8	5806,5	6365,6
15D	20	5722,7	5145,7	6231,1	5465,1	4954,5	4767,0	6272,2	6338,6	5348,5	5732,3
30D	5	20154,4	19179,4	19592,4	19521,9	19862,1	19398,8	19304,2	19419,7	19570,0	20201,6
30D	20	20196,3	19237,1	19670,9	19406,0	19811,9	19592,1	18728,8	19228,1	19859,6	19417,2
45D	5	42943,3	42439,7	42376,9	42582,9	43055,5	43745,1	41682,4	43156,3	43369,7	43803,2
45D	20	41225,0	40810,4	41342,7	40609,3	41551,1	42430,8	40287,6	41587,0	42322,8	41349,4

Limite Inferior do Custo da Solução Óptima

D.R.	N.T.C.	1	2	3	4	5	6	7	8	9	10
10S	5	980,4	860,7	935,4	973,2	939,3	940,8	928,6	849,9	819,6	917,1
10S	20	969,0	860,6	925,0	973,0	938,5	964,9	1044,4	850,0	820,6	918,6
15S	5	2808,1	2590,6	2573,1	2444,3	2487,8	2396,1	2554,5	2355,7	2471,7	2523,9
15S	20	2744,3	2521,8	2455,5	2482,4	2583,6	2461,1	2545,3	2357,2	2465,8	2458,9
30S	5	13792,4	14318,0	14502,3	14490,0	14447,8	14309,3	14323,9	13779,0	14728,7	13756,1
30S	20	13732,3	14243,3	14407,8	14410,7	14342,0	14226,2	14185,9	13690,5	14678,0	13696,1
45S	5	38206,5	38457,6	38939,8	38977,4	37830,6	37400,7	37765,4	38802,1	37767,3	39043,9
45S	20	37074,8	37142,2	37805,2	37736,5	36808,7	36167,5	36421,5	37315,7	36601,1	37785,0

D.R.	N.T.C.	1	2	3	4	5	6	7	8	9	10
10D	5	702,8	703,7	717,0	709,8	676,3	782,2	812,5	714,8	803,9	713,6
10D	20	701,7	740,5	716,5	644,7	675,7	840,7	816,7	717,6	790,7	705,5
15D	5	2011,1	2011,9	1750,3	2003,0	1876,8	1963,4	2072,6	2234,4	2035,5	1944,2
15D	20	1910,4	1914,8	1732,3	1986,3	1939,2	1959,6	2097,0	2188,2	2099,6	1902,2
30D	5	11710,5	10922,2	11394,3	11229,5	11607,7	11362,7	10921,2	11231,2	11268,4	11452,4
30D	20	11326,0	10573,3	11001,6	10851,0	11186,4	10956,8	10579,3	10902,1	10879,8	11064,5
45D	5	31629,8	31132,2	30869,7	31273,7	31430,9	32623,1	30406,8	31612,0	31948,9	31795,4
45D	20	29657,7	29264,9	29174,0	29228,8	29700,0	30721,8	28989,4	29862,5	30413,2	29943,7

Tempo de Cálculo (segundos)

D.R.	N.T.C.	1	2	3	4	5	6	7	8	9	10
10S	5	2,11	2,07	2,06	2,10	2,10	2,14	2,08	2,10	2,07	2,17
10S	20	6,59	6,56	6,58	6,55	6,65	6,51	6,56	6,53	6,62	6,67
15S	5	7,49	7,56	7,49	7,56	7,59	7,54	7,59	7,70	7,62	7,59
15S	20	23,89	23,88	23,86	24,47	24,36	24,43	24,13	24,58	24,41	24,17
30S	5	64,60	64,78	64,93	65,10	64,87	64,26	64,94	66,56	65,45	64,44
30S	20	211,30	211,11	211,03	211,63	211,58	211,38	211,80	211,82	211,19	211,51
45S	5	231,53	231,10	232,31	232,53	233,51	233,70	232,23	231,81	231,46	231,60
45S	20	705,50	705,63	704,66	707,65	706,15	706,99	706,65	705,96	709,40	706,37

D.R.	N.T.C.	1	2	3	4	5	6	7	8	9	10
10D	5	3,19	3,26	3,25	3,23	3,22	3,30	3,24	3,25	3,25	3,22
10D	20	10,60	10,54	13,09	10,59	10,61	10,54	10,55	10,55	10,66	10,59
15D	5	12,00	11,94	11,94	11,98	12,02	11,95	12,06	12,05	12,09	12,15
15D	20	37,97	38,04	38,25	38,65	38,99	39,53	40,01	39,38	38,46	38,54
30D	5	107,76	108,34	108,36	108,55	108,40	108,24	108,07	108,23	108,97	108,57
30D	20	337,10	336,56	336,60	336,88	337,02	336,98	336,93	336,85	336,83	336,71
45D	5	378,30	385,14	383,42	379,15	380,81	379,75	379,57	380,11	380,17	379,93
45D	20	1148,37	1147,93	1151,45	1152,26	1149,59	1148,79	1149,71	1150,48	1150,42	1148,87

ANEXO C

ALGORITMO *BRANCH-AND-BOUND*

O texto aqui apresentado baseia-se na descrição deste algoritmo feita em [Mital83]. Em teoria, o algoritmo *Branch-and-Bound* permite calcular a solução óptima de qualquer problema de Programação Inteira (o caso em que todas as variáveis são inteiras) ou Programação Inteira Mista (o caso em que algumas variáveis são inteiras e outras são reais) embora na prática, o tempo de cálculo para muitos problemas reais seja demasiado longo. Tal como o nome indica, este algoritmo consiste em aplicar alternadamente duas estratégias. Uma estratégia é decompor (*branch*) um problema em dois sub-problemas e a outra estratégia é resolver cada um dos sub-problemas de modo a calcular limites (*bounds*) apropriados para a função objectivo do problema original. No caso de problemas de minimização, o objectivo é calcular limites inferiores e no caso de problemas de maximização, o objectivo é calcular limites superiores.

Considere-se um problema de minimização com N variáveis das quais as variáveis x_i, $i = 1, 2, \ldots, R$, são inteiras e as variáveis x_i, $i = R+1, \ldots, N$, são reais. Um problema de Programação Inteira ($R = N$) é um caso particular de um problema de Programação Inteira Mista pelo que a descrição que se segue é genérica para os dois casos.

A relaxação linear do problema original consiste no problema em que a função objectivo e as restrições são iguais ao problema original e em que se consideram todas as

variáveis reais. O termo relaxação linear significa a eliminação das restrições que impõem que as variáveis sejam inteiras. A resolução da relaxação linear é muito mais simples do que a resolução do problema original. Designando a relaxação linear do problema original por Problema 1, a sua solução é um limite inferior da solução óptima do problema original.

Se as variáveis x_i, $i = 1, 2, \ldots, R$, da solução do Problema 1 forem inteiras (diz-se neste caso que o Problema 1 tem uma solução inteira), então está encontrada a solução do problema original. Se isso não acontecer, então pelo menos uma das variáveis x_i, $i = 1, 2, \ldots, R$, é não inteira. Seja x_p uma dessas variáveis, e seja $x_p = b$, o valor não inteiro dessa variável na solução do Problema 1 (considere $[b]$ como sendo o maior inteiro menor do que b). Com base no Problema 1, definem-se então dois sub-problemas: Problema 11 e Problema 12. O Problema 11 é definido pelo Problema 1 em que se acrescenta a restrição $x_p \leq [b]$ e o Problema 12 é definido pelo Problema 1 em que se acrescenta a restrição $x_p \geq [b]+1$. Neste caso, diz-se que o Problema 1 é o problema pai do Problema 11 e Problema 12. Esta operação de decomposição separa o espaço de soluções possíveis em dois conjuntos e a solução do problema original existe necessariamente num deles pois o espaço de soluções eliminado (todas as soluções tais que $[b] < x_p < [b]+1$) não exclui nenhuma solução do problema original.

Cada um dos dois problemas lineares, Problema 11 e Problema 12, é agora tratado individualmente tal como o descrito para o Problema 1. O algoritmo evolui pela decomposição de cada problema em dois novos problemas lineares e pelo cálculo dos limites dados pelas soluções de cada um dos novos problemas gerados. De notar que o valor da solução de um problema é sempre maior ou igual do que o valor da solução do seu problema pai. A um problema gerado deixa de se aplicar a operação de decomposição quando acontece uma de três condições possíveis:

(i) o problema tem uma solução inteira, ou seja, uma solução cujo valor das variáveis x_i, $i = 1, 2, \ldots, R$, é inteiro; neste caso, foi encontrada uma solução possível para o problema original e a operação de decomposição é desnecessária pois nunca mais encontrará soluções inteiras de valor menor do que a solução encontrada;

(ii) o problema tem uma solução (inteira ou não inteira) cujo valor é maior ou igual do que a melhor solução inteira encontrada até ao momento para o problema original;

(iii) o problema não tem solução.

Quando todos os problemas gerados estão em alguma das condições descritas o algoritmo termina. A solução do problema original é a solução inteira de menor valor de todos os problemas que terminaram pela condição (i). Se nenhum problema terminou nesta condição, então todos os problemas terminaram na condição (iii) e o problema original não tem solução.

Vejamos um exemplo que ilustra o algoritmo descrito. Considere-se o seguinte problema:

Minimizar: $-3x_1 - 4x_2$

Tal que: $2x_1 + 4x_2 \leq 13$

$-2x_1 + x_2 \leq 2$

$2x_1 + 2x_2 \geq 1$

$6x_1 - 4x_2 \leq 15$

$x_1, x_2 \geq 0$ e inteiros

Eliminando a condição que impõe as variáveis x_1 , x_2 serem inteiras, obtém-se um problema linear cuja solução é dada por $x_1 = 3.5$, $x_2 = 1.5$ e cujo valor da função objectivo é –16.5. Esta é a solução do Problema 1 na Figura C.1 e o valor da solução do problema é um limite inferior para a solução do problema original. Dado que a variável x_1 é inteira no problema original, geramos os problemas 11 e 12 impondo as restrições $x_1 \leq 3$ e $x_1 \geq 4$. O problema 12 não tem solução pelo que é eliminado pela condição (iii). O problema 11 não verifica nenhuma das condições anteriores e gera os problemas 111 e 112 por decomposição aplicada neste caso à variável x_2 (Figura C.1). O problema 111 gera uma solução inteira ($x_1 = 3$, $x_2 = 1$) pelo que este problema não é mais decomposto pela condição (i). O problema 112 é decomposto porque o seu limite –15.5 é menor do que a solução inteira até aí encontrada (-13 no problema 111) pelo que pode gerar uma solução inteira com um valor algures entre –15.5 e –13. A sua decomposição resulta num problema impossível (problema 1122) e no problema 1121 cujo limite é agora –15. O problema 1121 é finalmente decomposto no problema 11211 (que tem uma nova solução inteira) e no problema 11212. O problema 11211 termina pela condição (i) e o valor da nova solução inteira ($x_1 = 2$, $x_2 = 2$) é –14. O limite do problema 11212 é –13.5 pelo que este problema é eliminado pela condição (ii) ou seja, o seu limite é maior do que o valor de uma solução inteira entretanto encontrada. A solução para o problema original é dada pelo problema 11211 que é a menor solução inteira de todas as encontradas.

Note-se que o limite de cada um de problemas gerados é sempre maior ou igual do que o limite do problema pai. Note-se também que em todos os instantes, o menor valor dos limites dos problemas que estão em condições de serem decompostos é um limite para o problema original.

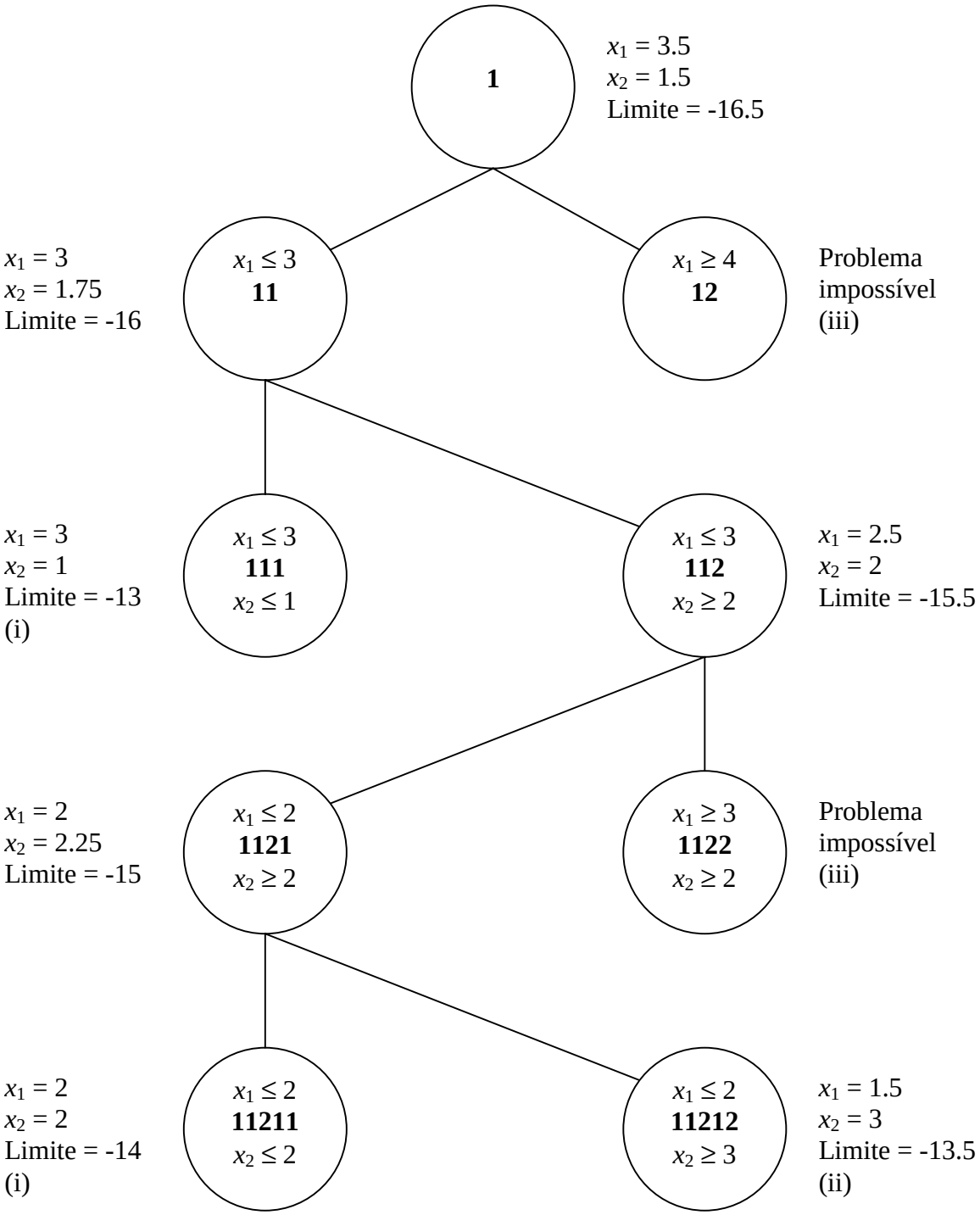


Figura C.1 : Ilustração do algoritmo de *branch and bound*

```
      Nodes                                         Cuts/
 Node  Left     Objective  IInf  Best Integer    Best Node    ItCnt

    0     0     1616.7782    14                 1616.7782      159
  100    98     6687.0969    17                 1627.4873     2846
  200   188     7844.4932    27                 1641.8359     6110
  300   252     2266.4667    17                 1739.3385     9700
*  363   198     5060.4000     0     5060.4000   1739.3385    10773
  400   233     3310.3566    26     5060.4000   1753.8190    12073
  500   333     2763.4553    21     5060.4000   1844.7267    14954
  600   431     3889.0949    15     5060.4000   1872.8646    17092
  700   528     3936.8238    27     5060.4000   1900.2539    19652
  800   628     3854.4460    20     5060.4000   1904.7431    22192
  900   722     3014.9230    17     5060.4000   1916.7131    25053
 1000   822     4244.3373    29     5060.4000   1938.4084    28172
Elapsed b&b time =   3.35 sec. (tree size =  0.42 MB)
 1100   920     3559.3986    11     5060.4000   1948.9554    30403
 1200  1015     2786.7921    29     5060.4000   1969.3917    32222
 1300  1113     4383.9618    23     5060.4000   1972.2794    35072
 1400  1206     2403.6301    14     5060.4000   1981.7726    38086
 1500  1303     3461.0657    29     5060.4000   1991.0393    41135
 1600  1397     2729.6676    31     5060.4000   2001.3802    43505
 1700  1494     3534.1418    25     5060.4000   2016.3489    46298
 1800  1590     2696.6417    20     5060.4000   2023.6108    48681
```

Figura C.2 : CPLEX: Início da resolução de um problema de Programação Inteira

```
 9800  4318     2499.0212     8     2635.2000   2477.4380   266849
 9900  4372     2516.4219    17     2635.2000   2477.6865   269365
10000  4430        cutoff              2635.2000   2478.3701   271858
Elapsed b&b time =  25.98 sec. (tree size =  2.48 MB)
10100  4481     2576.9732    16     2635.2000   2478.6094   273582
10200  4543     2539.8511    18     2635.2000   2478.9666   276070
10300  4587        cutoff              2635.2000   2479.5206   278309
10400  4648     2583.2774    14     2635.2000   2479.6259   280693
* 10412  3288     2590.5000     0     2590.5000   2479.6259   281003
* 10476  3127     2585.2000     0     2585.2000   2479.7663   282693
10500  3131     2499.2180    29     2585.2000   2479.8316   283164
10600  3148     2566.8968    19     2585.2000   2480.5380   285544
10700  3178     2583.5000    21     2585.2000   2481.0380   287916
10800  3214     2537.3264    16     2585.2000   2481.0380   290302
10900  3260     2558.3365    11     2585.2000   2481.0380   292525
11000  3316     2481.0380     8     2585.2000   2481.0380   295042
Elapsed b&b time =  29.06 sec. (tree size =  1.85 MB)
11100  3368     2528.8357    11     2585.2000   2481.0569   297527
11200  3409    infeasible           2585.2000   2481.4536   299832
11300  3437     2546.4500    22     2585.2000   2481.8382   302106
11400  3459        cutoff              2585.2000   2482.3623   304049
11500  3474     2483.3663     7     2585.2000   2482.7171   305506
11600  3518     2530.2363    24     2585.2000   2483.0864   307987
11700  3554     2512.1779    14     2585.2000   2483.2259   310220
11800  3601     2507.6632     1     2585.2000   2483.2259   312790
```

Figura C.3 : CPLEX: Continuação da resolução de um problema de Programação Inteira

```
54000   747     2524.9094    13     2527.5000   2524.9094   850794
Elapsed b&b time =  79.36 sec. (tree size =  0.61 MB)
54100   723     2527.0279    18     2527.5000   2525.0370   851438
54200   678        cutoff              2527.5000   2525.2254   852057
54300   618     2525.3176     6     2527.5000   2525.3176   852395
54400   596     2525.4309    11     2527.5000   2525.4309   853023
54500   575        cutoff              2527.5000   2525.5552   854095
54600   528        cutoff              2527.5000   2525.6413   854361
54700   475    infeasible           2527.5000   2525.7959   854982
54800   428     2525.9293    19     2527.5000   2525.9293   855518
54900   391    infeasible           2527.5000   2526.0934   856405
55000   343    infeasible           2527.5000   2526.2915   857084
Elapsed b&b time =  80.90 sec. (tree size =  0.35 MB)
55100   300        cutoff              2527.5000   2526.4496   857686
55200   262    infeasible           2527.5000   2526.5999   858490
55300   243     2526.7380     7     2527.5000   2526.7380   859079
55400   205        cutoff              2527.5000   2526.7839   859920
55500   152        cutoff              2527.5000   2526.9734   860603
55600   101     2527.1528    19     2527.5000   2527.1528   861123

Integer optimal solution (0.0001/0):  Objective =   2.5275000000e+003
Current MIP best bound =   2.5272501068e+003 (gap = 0.249893)
Solution time =   81.97 sec.  Iterations = 861697  Nodes = 55678 (73)
```

Figura C.4 : CPLEX: Conclusão da resolução de um problema de Programação Inteira

As figuras C.2, C.3 e C.4 ilustram a evolução do algoritmo de *branch and bound* quando executado pela ferramenta CPLEX [CPLEX98]. Esta ferramenta designa cada problema por nó (*node*). As grandezas mais significativas são as seguintes. A coluna *Node* indica quantos problemas foram processados até ao momento. Um problema processado é um problema que ou foi eliminado ou foi decomposto em dois novos problemas. A coluna *Nodes Left* indica quantos problemas foram gerados mas ainda não foram processados. A coluna *Best Integer* indica qual o menor valor de todas as soluções inteiras entretanto encontradas. Finalmente, a coluna *Cuts / Best Nodes* refere-se a dois tipos de informação. No exemplo apresentado nas figuras, esta coluna refere-se sempre ao menor valor de todos os problemas gerados mas não processados.

Quando o algoritmo é iniciado (Figura C.2), a coluna *Best Integer* está em branco até que seja encontrada a primeira solução inteira. O valor da coluna *Best Node* é, em cada instante, um limite inferior da solução do problema original e vai aumentando durante a execução do algoritmo. O algoritmo termina quando este limite inferior é igual ao valor da melhor solução inteira encontrada (Figura C.4), ou seja, quando o valor da coluna *Best Integer* iguala o valor da coluna *Best Node*. Note-se finalmente que quando o algoritmo encontra uma solução inteira melhor, o número de problemas não processados (coluna *Nodes Left*) diminui. Esta diminuição é maior ou menor consoante a melhoria do valor de *Best Integer*. Por exemplo, a Figura C.3 captura duas melhorias do valor de *Best Integer*. A primeira é muito mais significativa que a segunda pelo que na primeira, o número de nós não processados diminui muito mais. Esta redução acontece porque quando o limite superior (valor de *Best Integer*) desce de um valor *A* para um valor *B*, os problemas gerados mas não processados cujos valores estão entre *A* e *B* são automaticamente eliminados pela condição (ii).

Finalmente, uma referência à informação da coluna *Objective*. Esta coluna indica o valor da solução do último problema processado. No caso apresentado, esta informação é indicada de 100 em 100 problemas processados mais todas as vezes que encontra uma solução inteira de valor menor que as anteriores. Quando o problema processado é eliminado pela condição (iii), esta coluna indica *infeasible* e quando é pela condição (ii), esta coluna indica *cutoff*. Note-se que quando o algoritmo se aproxima do seu fim, o número de problemas que são eliminados torna-se maior como se pode verificar pelo número de vezes em que as indicações *infeasible* e *cutoff* aparecem nas 3 figuras.

LISTA DE ACRÓNIMOS

ATM	Asynchronous Tranfer Mode
ATM Forum	Organismo de normalização promovido pelos fabricantes de equipamentos
CBR	Constant Bit Rate na nomenclatura do ATM Forum (o mesmo que DBR)
CDV	Cell Delay Variation
CLR	Cell Loss Rate
CTD	Cell Transit Delay
DBR	Deterministic Bit Rate na nomenclatura do ITU-T (o mesmo que CBR)
DS1	Ligação a 1.5 Mbps da hierarquia PDH americana
E1	Ligação a 2 Mbps da hierarquia PDH europeia
E3	Ligação a 34 Mbps da hierarquia PDH europeia
IETF	Internet Engineering Task Force – forum de proposta, discussão e aprovação de normas para a Internet aberto a toda a comunidade científica
IP	Internet Protocol
ITU	International Telecommunications Union - organismo de normalização de âmbito mundial
ITU-T	Sector de Telecomunicações do ITU (antigo CCITT)
LAN	Local Area Network (Rede de Área Local)
MBS	Maximum Burst Size
MPLS	MultiProtocol Label Swithcing
PCR	Peak Cell Rate
PCPS	Primeiro a Chegar Primeiro a Sair (disciplina de filas de espera)
PDH	Plesiochronous Digital Hierarchy

RSVP	Resource ReSerVasion Protocol
SBR	Statistical Bit Rate na nomenclatura do ITU-T (o mesmo que VBR)
SCR	Sustainable Cell Rate
SDH	Synchronous Digital Hierarchy
STM-1	Ligação a 155 Mbps na hierarquia SDH
STM-4	Ligação a 622 Mbps na hierarquia SDH
VBR	Variable Bit rate na nomenclatura do ATM Forum (o mesmo que SBR)
VCC	Virtual Channel Connection
VCI	Virtual Channel Identifier (campo do cabeçalho de uma célula ATM)
VPC	Virtual Path Connection
VPI	Virtual Path Identifier (campo do cabeçalho de uma célula ATM)

REFERÊNCIAS

[Ahn94] S. Ahn, R. Tsang, S. Tong, D. Du, "Virtual Path Layout Design on ATM Networks", *Proc. IEEE INFOCOM'94*, Toronto, Canadá, 1994, pp. 192-200

[Al-Rumaih96] A. Al-Rumaih, R. Ahmed, S. Bakry, A. Al-Dhelaan, "A Methodology for Network Topology Design with Link and Node Failures Tolerance", *ACM Int. Journal of Network Management*, Vol. 6, No. 1, pp. 42-63, 1996

[Anerousis96] N. Anerousis, A. Lazar, "An Architecture for Managing Virtual Circuit and Virtual Path Services on ATM Networks", *Journal of Network and System Management*, Vol. 4, No. 4, Dezembro 1996

[Anerousis98] N. Anerousis, A. Lazar, "Virtual Path Control for ATM Networks with Call Level Quality of Service Guarantees", *IEEE/ACM Transactions on Networking*, Vol. 6, No. 2, Abril 1998

[Anick82] D. Anick, D. Mitra, M. Sondi "Stochastic theory of a data handling system with multiple sources", *Bell Syst. Tech. J.*, Vol. 61, No. 8, pp. 1871-1894, Outubro 1982

[ATMForum99] ATM Forum, "Inverse Multiplexing for ATM (IMA) Specification – Version 1.1", Março 1999

[Balakrishnan89] A. Balakrishnan, T. Magnanti, R. Wong, "A Dual-Ascent Procedure for Large Scale Uncapacitated Network Design", *Operations Research*, Vol. 37, pp. 716-740, 1989

[Beasley93] J. Beasley, "Lagrangean Relaxation", *Modern Heuristic Techniques for Combinatorial Problem*, Colin Reeves (editor), Blacwell Scientific, 1993

[Bienstock96] D. Bienstock, O. Günlük, "Capacitated Network Design – Polyhedral Structure and Computation", *INFORMS Journal on Computing*, Vol.8, No. 3, 1996, pp. 243-257

[CCITT92a] CCITT Rec. I.311, "B-ISDN General Network Aspects", Junho 1992

[CCITT92b] CCITT Rec. I.361, "B-ISDN ATM Layer Specification", Junho 1992

[Cheng94] K. Cheng, F. Lin, "On the Joint Virtual Path Assignment and Virtual Circuit Routing Problem in ATM Networks", *Proc. IEEE GLOBECOM'94*, San Francisco, EUA, pag. 436-441, 1994

[Chlamtac93] I. Chlamtac, A. Faragó, T. Zhang, "How to Establish and Utilize Virtual Paths in ATM Networks", *Proc. IEEE ICC'93*, Geneve, Suiça, 1993, pp. 1368-1372

[Chlamtac94] I. Chlamtac, A. Faragó, T. Zhang, "Optimizing the System of Virtual Paths", *IEEE/ACM Transactions on Networking*, Vol. 2, No. 6, Dezembro 1994, pp. 581-587

[Cidon94] I. Cidon, O. Gerstel, S. Zaks, "A Scalable Approach to Routing in ATM Networks", *Proc. 8th Int. Workshop on Distributed Algorithms*, pp. 208-222, Outubro 1994

[Comm95] IEEE Communications Magazine, Número especial sobre "Dynamic Routing in Telecommunications Networks", Vol. 33, No. 7, Julho 1995

[Comm99] IEEE Communications Magazine, Número especial sobre "Multiprotocol Label Switching", Vol. 37, No. 12, Dezembro 1999, pp. 36-68

[COST94] M. Ritter, P. Tran-Gia (editors), "Multi-Rate Models for Dimensioning and Performance Evaluation of ATM Networks", *COST 242 Interim Report*, Junho 1994

[CPLEX98] CPLEX Software Optimisation Package, ILOG Inc. Versão 6.0, 1998

[Dziong97] Z. Dziong, "ATM Network Resource Management", McGrow-Hill, 1997

[Elawlid93] A. Elawlid, D. Mitra, "Effective Bandwidth of General Markovian Traffic Sources and Admission Control of High Speed Networks", *IEEE/ACM Transactions on Networking*, Vol. 1, No. 3, Junho 1993.

[Evans91] S. Evans, "Optimal Bandwidth Management and Capacity Provision in a Broadband Network using Virtual Paths", *Performance Evaluation*, Vol. 13, No. 1, pp. 27-43, 1991

[Faragó95] A. Faragó, S. Blaabjerg, L. Ast, G. Gordos, T. Henk, "A New Degree of Freedom in ATM Network Dimensioning: Optimizing the Logical Configuration", *IEEE Journal on Selected Areas in Communications*, Vol. 13, No. 17, Setembro 1995, pp. 1199-1206.

[Frisen96] V. Frisen, J. Harms, J. Wong, "Resource Management with Virtual Paths in ATM Networks", *IEEE Network*, Vol. 10, No. 5, pp. 10-20, Set./Out. 1996

[Gavish89] B. Gavish, I. Neuman, ""A System for Routing and Capacity Assignment in Computer Communication Networks", *IEEE Transactions on Communications*, Vol. 37, pp. 360-366, 1989

[Gavish90] B. Gavish, A. Altinkemer, "Backbone Network Design Tools with Economic Tradeoffs", *ORSA Journal of Computing*, Vol. 2, pp. 236-252, 1990

[Geoffrion74] A. M. Geoffrion, ""Lagrangean Relaxation and its Uses in Integer Programming ", *Mathematical Programming Study*, Vol. 2, pp. 82-114, 1974

[Gerla89] M. Gerla, J. Monteiro, R. Pazos, "Topology Design and Bandwidth Allocation in ATM Nets", *IEEE J. Sel. Areas in Communications*, Vol. 7, No. 8, pp. 1253-1262, Outubro 1989

[Gersht94] A. Gersht, A. Shulman, "Optimal Dynamic Virtual Path Bandwidht Allocation and Restoration in ATM Networks", *Proc. IEEE GLOBECOM'94*, San Francisco, EUA, pp. 770-776, 1994

[Gerstel96] O. Gerstel, I. Cidon, S. Zaks, "The Layout of Virtual Paths in ATM Networks", *IEEE/ACM Transactions on Networking*, Vol. 4, No. 6, pp. 873-884, Dezembro 1996

[Gibbens91] R. Gibbens, P. Hunt, "Effective bandwidths for the multi-type UAS channel", *Queueing Systems*, Vol. 9, pp. 17-28, 1991

[Gibbens95] R. Gibbens, F. Kelly, "Network Programming Methods for Loss Networks", *IEEE J. Sel. Areas Comm.*, Vol. 13, No. 7, pp. 1189-1198, 1995

[Girard98] A. Girard, B. Sansó, "Multicommodity Flow Models, Failure Propagation, and Reliable Loss Netowrk Design", *IEEE/ACM Transactions on Networking*, Vol. 6, No. 1, pp. 82-93, Fevereiro 1998

[Gupta92] S. Gupta, K. Ross, "Routing in Virtual Path Based ATM Networks", *Proc. IEEE GLOBECOM'92*, pp. 571-575, 1992

[Held70] M. Held, R. Karp, "The travelling-salesman problem and minimum spanning trees", *Operations Research*, 18, pp. 1138-1162, 1970

[Held71] M. Held, R. Karp, "The travelling-salesman problem and minimum spanning trees: Part II", *Operations Research*, 1, pp. 6-25, 1971

[Held74] M. Held, P. Wolfe, H. Crowder, "Validation of subgradient optimisation". *Mathematical Programming*, vol. 6, pp. 62-88, 1974

[Hwang94] R. Hwang, J. Kurose, D. Towsley, "MDP Routing in ATM Networks Using Virtual Path Concept", *Proc. IEEE INFOCOM'94*, Toronto, Canadá, pp. 1509-1517, 1994

[Hwang95] R. Hwang, "LLR Routing in Homogeneous VP-based ATM Networks", *Proc. IEEE INFOCOM'95*, Boston, EUA, pp. 587-593, 1995

[IETF97] R. Braden, L. Zhang, S. Berson, S. Herzog, S. Jamin, "Resource reSerVation Protocol (RSVP) – Version 1 Functional Specification", RFC 2205, *IETF Network Working Group*, Setembro 1997

[IETF99] E. Rosen, A. Viswanathan, R. Callon, "Multiprotocol Label Switching Architecture", Internet draft draft-ietf-mpls-arch-05.txt, Abril 1999

[Iraschko98] R. Iraschko, M. MacGregor, W. Grover, "Optimal Capacity Placement for Path Restoration in STM and ATM Mesh-Survivable Networks", *IEEE/ACM Transactions on Networking*, Vol. 6, No. 3, pp. 325-335, Junho 1998

[Jagerman84] D. Jagerman, "Methods in Traffic Calculations", *AT&T Bell Laboratories Technical Journal*, Vol. 63, No. 7, pp. 1283-1310, Setembro 1984

[Kelly86] F. Kelly, "Blocking Probabilities in Large Circuit-Switched Networks", *Advances in Applied Probability*, Vol. 18, pp. 473-505, 1986

[Kelly91] F. Kelly, "Effective Bandwidths in Multi-Class Queues", *Queueing Systems*, Vol. 9, pp. 5-16, 1991

[Kesidis93] G. Kesidis, J. Walrand, "Effective Bandwidths for Multiclass Markov Fluids and Other ATM Sources", *IEEE/ACM Transactions on Networking*, Vol. 1, No. 4, pp. 424-428, Agosto 1993

[Kim95] S. Kim, "An Optimal Establishment of Virtual Path Connections for ATM Networks", *Proc. IEEE INFOCOM'95*, Boston, EUA, pp. 72-79, 1995

[Kim96] S. Kim, M. Kim, D. Jang, "Optimization Models and Algorithms for Planning ATM Networks", *Proc. IEEE GLOBECOM'96*, 1996

[LeBlanc89] L. LeBlanc, R. Simmons, "Continuous Models for Capacity Design of Large Packet Switched Telecommunications Networks", *ORSA Journal of Computing*, Vol. 1, pp.271-286, 1989

[Lin93] F. Lin, K. Cheng, "Virtual Path Assignment and Virtual Circuit Routing in ATM Networks, *Proc. IEEE GLOBECOM'93*, Houston, EUA, pp. 436-441, 1993

[Magnanti95] T. Magnanti, P. Mirchandani, R. Vachani, "Modeling and Solving the Two Facility Capacitated Network Loading Problem", *Operations Research*, Vol. 43, No. 1, pp.142-156, 1995

[Martins97] E. Martins, M. Pascoal, J. Santos, "A New Algorithm for Ranking Loopless Paths", *Research Report*, CISUC, Universidade de Coimbra, 1997

[Maruyama77] K. Maruyama, K. Fratta, D. Tang, "Heuristic Design Algorithm for Computer Communication Networks with Different Classes of Costumers", *IBM Journal of Research and Development*, pp. 360-369, 1977

[Medhi94] D. Medhi, "Models foe Network Design and Management of ATM Networks based on Virtual Paths", *Computer Networks and ISDN Systems*, Agosto 1994

[Medhi95] D. Medhi, "Multi-Hour, Multi-Traffic Class Network Design for Virtual Path-Based Dynamically Reconfigurable Wide-Area ATM Networks", *IEEE/ACM Transactions on Networking*, Vol. 3, No. 6, Dezembro 1995

[Medhi97] D. Medhi, S. Guptan, "Network Dimensioning and Performance of Multi-Service Multi-Rate Loss Networks with Dynamic Routing", *IEEE/ACM Transactions on Networking*, Vol. 5, No. 6, pp. 944-957, Dezembro 1997

[Mital83] K. V. Mital, "Optimisation Methods", Wiley Eastern Limited, 2ª Edição, 1983

[Mitra96] D. Mitra, J. Morrison, K. Ramakrishnan, "ATM Network Design and Optimization: a Multirate Loss Network Framework" *IEEE/ACM Transactions on Networking*, Vol. 4, No. 4, pp. 531-543, 1996

[Murakami94] K. Murakami, H. Kim, "Near-Optimal Virtual Path Routing for Survivable ATM Networks", *Proc. IEEE INFOCOM'94*, Toronto, Canadá, pp. 208-215, 1994

[Ng87] T. Ng, D. Hoang, "Joint Optimization of Capacity and Flow Assignment in a Packet-Switched Communication Network", *IEEE Trans. On Computing*, Vol. 35, pp.202-209, 1987

[Onvural95] R. Onvural, "Asynchronous Transfer Mode Networks: Performance Issues", 2ª Edição, Artech House, 1995

[Ross95] K. Ross, "Multiservice Loss Models for Broadband Telecommunication Networks", Springer-Verlag, 1995

[Sato91] Y. Sato, K. Sato, "Virtual Path and Link Capacity Design for ATM Networks", *IEEE Journal on Selected Areas in Communications*, Vol. 9, No. 1, January 1991, pp.104-111

[Sethi95] A. Sethi, "A Model for Virtual Tree Bandwidth Allocation in ATM Networks, *Proc. IEEE INFOCOM'95*, Boston, EUA, 1995, pp. 1222-1228

[Sousa99] A. de Sousa, R. Valadas, L. Cardoso, A. Duarte, "ATM Network Dimensioning for Mixed Symmetrical and Asymmetrical Services with Dynamic Reconfiguration in a Multi-Network Provider Environment", *Proc. Third IFIP Workshop on Traffic Management and Design of ATM Networks*, Londres, Reino Unido, pp. 5/1-5/15, 26 - 27 Abril, 1999

[Sousa01] A. de Sousa, R. Valadas, L. Gouveia, "ATM Network Dimensioning using 2-Layer Hierarchical Virtual Path Layouts", *Proc. 9th International Conf. on Telecommunications Systems*, Dallas, Texas, USA, 15-18 Março, 2001

[Valadas98] R. Valadas, "Dimensioning and Resource Management of ATM Networks", *Proc. 7th IFIP/ICCC Conference on Information Networks and Data Communications (INDC'98)*, Aveiro, Portugal, Junho 15-17, 1998

[Veciana94] G. de Veciana, C. Courcoubetis, J. Walrand, “Decoupling Bandwidths for Networks: A Decomposition Approach to Resource Management”, *Proc. IEEE INFOCOM'94*, Toronto, Canadá, pp. 466-473, Maio 1994

[Veitch97] P. Veitch, D. Johnson, “ATM Network Resilience”, *IEEE Network*, Vol. 11, No. 5, pp. 26-33, Set./Out. 1997

[Whitt85] W. Whitt, “Blocking when Service is Required from Several Facilities Simultaneously”, *AT&T Technical Journal*, 63, pp. 1807-1856, 1985

[Xiao99] X. Xiao, L. Ni, “Internet QoS: A Big Picture”, *IEEE Network*, pp. 8-18, Março/Abril 1999

[Xiong99] Y. Xiong, L. Mason, “Restoration Strategies and Spare Capacity Requirements in Self-Healing ATM Networks”, *IEEE/ACM Transactions on Networking*, Vol. 7, No. 1, pp. 98-110, 1999

[Yang95] T. Yang, D. Tsang, “A Novel Approach to Estimating the Cell Loss Probability in an ATM Multiplexer Loaded with Homogeneous On-Off Sources”, *IEEE Transactions on Communications*, Vol. 43, pp. 117-126, Janeiro 1995

[Zhang93] L. Zhang, S. Deering, D. Estrin, S. Shemker, D. Zappala, “RSVP: A New Resource ReSerVation Protocol”, *IEEE Network*, pp. 8-18, Setembro 1993