MASTER

Multipoint connection management in ATM networks

van Berchum, W.

Award date:
1995

Eindhoven University of Technology
Faculty of Electrical Engineering
Digital Information Systems Group
Eindhoven

PTT Research
Dr Neher Laboratorium
Leidschendam

Multipoint connection management in ATM networks

W. van Berchum

Master of Science Thesis Report
Leidschendam, March 1995

Coaches
prof. ir. J. de Stigter
dipl. inf. A.R. von Stieglitz
ir. M.J.G. Dirksen

Management summary

Nowadays public telecommunication networks are exclusively optimized for support of either interactive services or distributive services. These separate networks use different transmission and switching techniques and are managed and controlled in distinct ways. A typical example of a network for interactive services is the telephony network. Networks providing unidirectional distributive services, like broadcast TV and radio, are for example the cable TV networks. A new transmission and switching technique, the Asynchronous Transfer Mode (ATM), offers the capabilities to integrate both the interactive and distributive services. In addition to interactive user-to-user services and one-to-many distributive services, also new interactive multipoint-to-multipoint services can be cost effectively provided within the same integrated concept. Telecommunication services with more than two participants are called multipoint services. Examples of such services are videoconferencing, tele-education, closed user group multicast (remote symposium), information retrieval services or joint CAD/CAM applications. These multipoint services will embody a major part of future telecommunication services.

Integration of interactive point-point services, interactive multipoint services and distributive multipoint services puts additional requirements onto the network and the network management and control functions. The transmission and switching requirements can be met by using ATM as transmission technique. The management and control functionality is to be implemented in new information management architectures like TINA (Telecommunication Information Network Architecture) or using B-ISDN Signalling and Management functions. Since research on multipoint routing has been limited, this report specifies the multipoint service requirements for connection routing and develops solutions for several multipoint routing problems.

Integration of new multipoint network functionality in the public network constitutes a new challenge for telecom operators. The Asynchronous Transfer Mode provides a main part of the required multipoint capabilities and becomes therefore of strategic importance in realizing the new multipoint opportunities. However, another main part of the multipoint requirements involves network management and control functions, which are addressed by the developed solutions for multipoint routing. Efficient and smart usage of network resources will determine competitive advantages against other network providers. This is even more important when the required network resources of each multipoint service are taken into account together with the level of multipoint service popularity.

The multipoint routing strategies developed in this report are based on service specific requirements and general network management and control objectives. The variety in multipoint service characteristics necessitates implementation of two distinct routing strategies: state-dependent and time-dependent routing. State-dependent routing adapts to actual traffic fluctuations and time-dependent routing adjusts to long term recurrent patterns in the network load. Multipoint routing requires special multipoint routing algorithms. Available algorithms are analyzed and judged on their appropriateness for the set up of multipoint connections in state-dependent and time-dependent routing strategies. The routing algorithms are either controlled centrally or controlled distributively. Time-dependent routing is performed most efficiently with centralized control algorithms. State-dependent routing requires a distributed approach in order to keep network control overhead low. Suitable algorithms are extended with dynamic features and a method is developed which provides an integrated routing approach for different multipoint services.

Technical summary

New telecommunication techniques like the Asynchronous Transfer Mode provide new multipoint capabilities which can be used to integrate new multipoint services in a telecommunication network. An extensive study of the available literature on multipoint services made clear that multipoint routing forms the main problem in multipoint connection management. This is caused by the complexity of the problem of finding an optimal multipoint routing tree in a network, the diverse service specific multipoint requirements and the constraints put on multipoint routing by general network resource management objectives. The main objective of this report is to deduce those requirements and constraints in order to design multipoint connection management functionality for all types of multipoint services.

The investigation on multipoint routing starts with the analysis of connection requirements for multipoint services. In order to deduce those connection requirements in a structured way, telecommunication services are first classified according to service specific attributes which are of main importance for connection management. The result is a class of interactive point-to-point services, a class of interactive conference services and a class of unidirectional distributive services. The main distinctions between conference services and distributive services apply to the number of participants and connection establishment. Conference services have a limited number of participants, distributive services have wide range of possible numbers of participants. Conference services require complete multipoint connections to be set up on demand while distributive services require (semi)-permanent connections. A basic requirement of both multipoint classes is that connections must adapt to dynamics in user groups caused by user join and leave events.

Connection set up requires reservation and allocation of network resources along the routes computed by routing procedures. Thus, these routing procedures play a major role in the network resource management. The general network resource management objectives include optimal utilization of network resources, avoidance of congestion situations, quality of service guarantees and low control overhead.

Both the service specific multipoint requirements and the network resource management objectives can be met by implementation of multipoint routing algorithms using routing strategies which adapt to traffic fluctuations in the network. State-dependent routing strategies adapt to short term variations in network load. Time-dependent routing strategies adapt to long term, recurrent patterns of traffic fluctuations. The variety in multipoint service characteristics necessitates use of both routing strategies.

The solutions developed in this report for the multipoint routing problem are based on implementation of routing algorithms in both routing strategies. A main conclusion is that a centralized control algorithm is needed for time-dependent routing and a distributed control algorithm is needed for state-dependent routing. Available algorithms from the literature are described and judged on their suitability for both routing strategies. It is also concluded that none of the available algorithms satisfies all service requirements and network requirements. Therefore, existing algorithms are extended with dynamic properties needed for adaption to user group dynamics. In addition, the asymmetric character of multipoint connections is not addressed by the available algorithms. The developed multipoint routing solution provides a link balancing mechanism and an integrated approach for unidirectional and bidirectional multipoint connections. The investigation on multipoint connections terminates with an analysis of the relationship between the proposed time-dependent and state-dependent routing algorithms and the network resource management.

Contents

1 Introduction

1.1 New capabilities for multipoint services

Availability of information has always been essential for the development of social and business activities. Telecommunications are in turn essential for wide spread availability of information. Growing demands for information in social and business life push technological developments in the area of telecommunications.
As telecommunication technologies emerge, new possibilities for telecommunication services are created. Traditionally, telecommunication networks provide either interactive services, like telephony, or broadcast services, like the distribution of television signals. The former yields bidirectional communication between two particular customers during an arbitrary time interval. The latter provides unidirectional communication from one source to all attached customers in a permanent way. These two extremities form the boundaries of the range of the so called multipoint services. In multipoint services an arbitrary number of participants communicate during any time interval. Examples are broadcast and multicast services for distribution of video and audio information, videoconferencing or tele-education. An example of a many-to-one contributive service is telemetry. New possibilities provided by emerging telecommunication technologies should be employed advantageously by integrating this range of new multipoint services in the public network.

One of the main objectives in the design of telecommunication networks is to integrate services with distinct characteristics within one technological concept. For efficiency reasons, the range of services to be supported by a new network technology should be as broad as possible. Currently, a new information transport technique is developed and standardized: the Asynchronous Transfer Mode (ATM). This new technique enables the desired integration of a wide range of services in one network.
Figure 1 illustrates the objective of realizing distinct telecommunication services using a universal platform of underlying transport techniques. The Asynchronous Transfer Mode offers this possibility by supporting connections for both point-to-point and multipoint telecommunication services, transporting any type of information (e.g. audio, video or data).

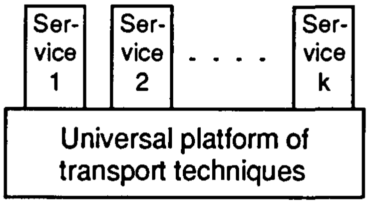


***Figure 1** Distinct services built on a universal platform of techniques*

Multipoint services distinguish themselves from point-to-point services by complex interconnection patterns, an arbitrary number of participants, distinct traffic flow directions, distinct service holding times etc. Thusfar, research in the field of multipoint services has been limited. Therefore, investigation on multipoint services and their network requirements is needed in order to integrate multipoint techniques in ATM and in the network management.

1.2 Definition of the problem

Connections through a network form the basis of every telecommunication service. Therefore, the main problem of this work on multipoint services is the identification and integrated design of multipoint connection management functionality for all types of multipoint services. Since connection routing forms a major function of the connection management, the multipoint routing problem must be solved. The solutions to be developed must satisfy service specific connection requirements and generally holding network resource management objectives. The analysis of these service specific connection requirements and network management objectives constitutes thus an integral part of the investigation on multipoint connections.

1.3 Readers guide

Chapter 2 starts with the analysis of all types of multipoint services and classifies them according to service characteristics which are relevant for connection management in general. These characteristics include communication symmetry, communication configuration, establishment of communication, dynamics and the number of service participants. Based on the resulting service requirements, chapter three develops ATM solutions for multipoint connections using generic network architectures. Chapter three also identifies general network resource management objectives and investigates how routing strategies can achieve these objectives.
Routing of multipoint connections forms the subject of chapter four. Available algorithms from literature are described and judged on their appropriateness for the distinct multipoint services and their applicability in the several routing strategies developed in chapter three. Suitable algorithms are extended where necessary and their consequences for the network control are described. Chapter 5 describes how connection management for all types of multipoint services can be integrated in one multipoint routing approach and how ATM network resource control functions are affected by this routing approach. Chapter 6 will eventually complete this investigation on multipoint services with conclusions and recommendations.

2 Multipoint services

This chapter deduces the network requirements of distinct multipoint services with respect to connection management. Main distinctions in characterizing service attributes lead to a general classification of telecommunication services. The analysis will focus on distinctions in connection requirements of two distinct multipoint service classes and will also recognize common requirements of services within the multipoint service classes.

2.1 Characterization of point-multipoint services

2.1.1 Attributes of telecommunication services

New technological developments in the area of telecommunication widen the range of possible services to be offered by a telecommunication network. A wide range of possible services necessitates characterization and classification of these services in order to determine technical requirements onto a network in a structured way. Arrangement of services in distinct classes according to specific characteristics ameliorates an integrated approach for the design of transport techniques supporting a complete class of services. Figure 2 illustrates the classification of services to be supported by a universal platform of techniques. These techniques provide ways to realize connections through the network which transport information entities containing user information and network control information.

Generally, telecommunication services can be characterized by describing its attributes and assigning values to these attributes. Distinct service classes can then be formed by combining services which have certain attributes and values in common. The attributes which are important for the deduction of requirements on connection management are (see [I140]):

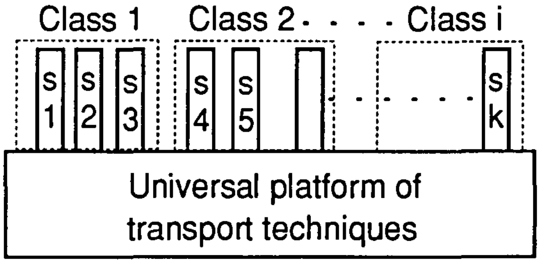


Figure 2 Classification of services

- establishment of communication
- symmetry of communication
- communication configuration
- user information type
- operational and commercial aspects
- service interworking
- quality of service

These attributes will be described first, then they will be evaluated in order to determine how they can be used to classify telecommunication services.

Establishment of communication

This service attribute describes the time characteristics of a communication session (*call*). Establishment of communication is determined by the start, duration and termination of the session. The three possible attribute values are:

- permanent
- semi-permanent
- on demand

In case of permanent establishment, the user is serviced by a permanent information stream, such as a 24-hour broadcast radio channel. A semi-permanent communication session has a predetermined time interval; start and termination times are known. Examples of semi-permanent services are non-continuous broadcast TV channels and church telephone services. The final value, on demand communication establishment, characterizes services with random start and termination times. In this case, communication establishment depends on user's behaviour (service request and termination) and on the speed of connection set up. An example is the well-known telephony service.
A subattribute of communication establishment is the *session holding time* (SHT). For example, TV broadcast sessions are characterized by long holding times, electronic mail services illustrate the adverse.
Another subattribute is defined by the service specific *time distribution*. The time distribution attribute indicates how frequently a given service is used or requested using a time scale based on days, weeks or even years. For example, the current telephony service is used most frequently at working hours (and New Years Day) while television entertainment services are used most frequently in the evenings.

Symmetry of communication

The symmetry of communication describes the relationship of information flow between two or more customers (also referred to as *access points*) involved in a communication. Possible values of the symmetry attribute are:

- unidirectional
- bidirectional symmetric
- bidirectional asymmetric

An example of unidirectional communication is the present television broadcast service. The customer is serviced by video signals transmitted through a coax cable network or via the ether while no communication in the backward direction takes place. Bidirectional symmetric communication can be compared with the plain old telephony service, where the communication path consists of two equally capacitated information streams between two users. A current example of bidirectional asymmetric communication is the File Transfer Protocol (FTP) service. The FTP user sends service requests (limited in transport capacity requirements) to an information server and is then serviced by a much more capacity demanding information stream providing the desired data.

Communication configuration

The communication configuration describes the spatial arrangement for transferring information between two or more access points. Possible values are:

- point-to-point
- point-to-multipoint
- multipoint-to-multipoint

Point-to-point (p-p) communication is rather straightforward. There are two access points involved in the communication session. Examples are a telephony call or a video on demand session, where single users are serviced by a video server with individual video information streams. Point-to-multipoint (p-mp) communication refers to communication between a single source/common destination access point and more than one other access points. In case of unidirectional point-to-multipoint communication, *multicast* and *broadcast* services can be distinguished. In the case of *multicast*, the number of destinations is limited and specified, while in the case of *broadcast* the number destinations is unlimited and unspecified. Multipoint-to-point communication is when the communication is unidirectional in the backward direction in a p-mp configuration. Multipoint-to-point characterizes services where information from a number of sources is collected by a joint destination. An example could be interactive television, where users respond individually to questions asked during a TV program, e.g. a quiz. Multipoint-multipoint (mp-mp) communication configurations are defined by bidirectional symmetric or asymmetric communication from multiple sources to multiple destinations (e.g. videoconferencing or chat-box).
The spatial arrangement for transferring information is yet characterized here by three additional subattributes. The first subattribute of the communication configuration is the *number of participants* in a communication session. The number of participants can be enormous in case of multicast or broadcast services, but will be limited in case of interactive services like telephony and videoconferencing. In addition, the number of participants can be limited in case of *closed user groups*. A closed user group is defined by a limited number of specified users who are licensed to participate in for that user group specific sessions of a specific service. An example is Church TV, where only church-members are licensed to participate in multicast sessions of divine services.
Another subattribute of communication configuration is formed by the *configuration dynamics*. The communication configuration can either be static throughout the whole session or non-static. In the latter case, the communication configuration changes while the session is in progress. This can be caused by users who either join or leave an existing session.

The final subattribute is the *geographical reach* of the communication configuration. This subattribute is especially important for broadcast and multicast services. In case of broadcast services the geographical reach determines the number of access points involved. In case of multicast services, the geographical reach is a measure of required network resources which provide the actual information flows.

User information type
This attribute characterizes a service by its required, service specific traffic type. Possible values are:

- speech
- video
- audio (sound)
- text
- computer data

This list is not exhaustive. Combinations of these information types are also possible. Within one information type, distinctions can be made between several (standardized) levels of qualities, e.g. picture quality in case of video.

Quality of Service

This attribute is described by a group of specific sub-attributes which apply for example to information loss, information mutilation, transfer delay or call set-up times.

Operational and commercial aspects

Several methods can be used in order to charge users of network services. They are represented by the following values of the subattribute *charging*:

- subscription based
- session based
- sponsor based

Subscription based charging requires a subscription contract between service provider and customer providing the customer the possibility to use a particular service for a particular period of time. The customer pays for the availability of the given service and the actual service usage does not influence the tariffs. The tariffs can be based on required information transfer rates and required quality of service.

Session based charging (also called 'pay per view' in case of television services) refers to payment per communication session. In this case, also the session holding time affects the revenues to be payed by service users.

Sponsor based charging refers to the situation where customers are not charged anyway because the service is sponsored by others. An example could be education broadcast television sponsored by the government.

Service interworking

This service attribute is especially important for services which differ only in a limited number of attributes from other services. An example is normal quality television compared with High Definition Television (HDTV).

General classification of services

Generally, the real-time transport of information through a network requires reservation of network resources in order to provide the required connections between communicating entities. It is the task of connection management to provide ways to set up and maintain these connections such that both the information transfer rate requirements and the quality of service requirements are met. However, in a telecommunication network in operation, network resources are limited available. Therefore, the management of connections plays a major role in the overall management of available network resources. In order to perform the task of resource management, it must be known by connection set up procedures how much network resources are required and where they are required in the network. Three service attributes mainly determine these network resource requirements, they are

- communication configuration
- communication symmetry
- information transfer rate.

The values of the first two attributes constitute three distinct service classes which are of major importance for resource management and thus for connection management. The three classes and their relations with the attribute values are illustrated in figure 2. The first class is the class of *interactive* services. This class consists of point-to-point services where two access points are involved in a non-unidirectional communication configuration. The class of *distributive* services contains unidirectional point-multipoint services (broadcast and multicast). The class of *conference* services consists of services which have both interactive and distributive aspects.

More than two access points are involved in an interactive, thus bidirectional, communication configuration. Examples are remote conferencing and tele-education services.
Distinct values of the information transfer rate attribute do not contribute to this classification, nevertheless, they are relevant to the connection management. Therefore, information transfer rates of distinct service types will be investigated in section 2.3.
Section 2.1.2 will go further into the class of distributive services. A number of distributive services are defined by assigning values to several attributes. The same is done for the class of conference services in section 2.1.3.

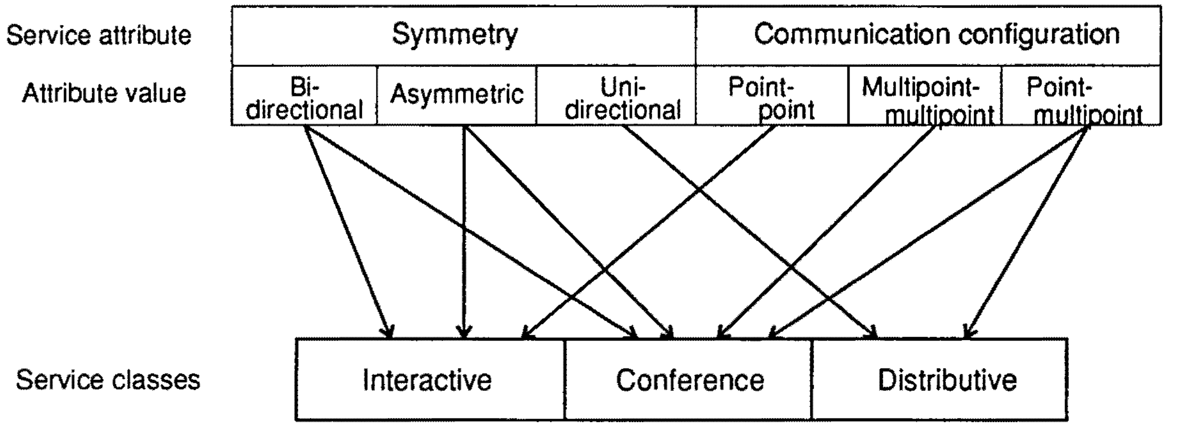


***Figure 3** General classification of telecommunication services*

2.1.2 Distributive services

Services are defined by assigning values to the attributes and subattributes as described in the previous section. Considering the class of distributive services, it was concluded that this class consists of unidirectional broadcast and multicast services. Broadcast services are characterized by unidirectional communication from one source to an unlimited number of unspecified destinations. Multicast services differ from broadcast services by their limited number of specified destinations. This section analyzes other distinctions in possible values of distributive service attributes.

Similarities in possible values of broadcast and multicast services attributes involve the information type attribute and the time distribution subattribute. Considering the information type, both broadcast and multicast services can be used for distribution of

- video information (TV programs)
- audio information (radio programs)
- files, data for graphics or text (information retrieval services).

Video information distribution services will mainly be used for entertainment purposes, just like television services today. Therefore, mass distribution of video using broadcast or multicast services will mainly be concentrated in spare time intervals, i.e. evenings and weekends. This is not the case for information retrieval services, because they are not predetermined entertainment services. Listening to radio programs is on the average not a main activity of the service user, therefore no assumptions can be made concerning the time distribution.
The non-interactive and unidirectional communication in broadcast and multicast services implicates that delay of transported information is of less importance than in case of interactive services.

Distinctions in possible values for broadcast and multicast services will be analyzed per attribute in the following.

Communication configuration

According to the definition, broadcast services are characterized by an unlimited number of participants. The only limitation on the number of participants is therefore the geographical reach of the broadcast service. It is evident that closed user groups cannot be implemented in broadcast services because destination endpoints are not specified. This is in contrast with multicast services, where closed user groups can be defined with specified destination endpoints. Generally, the number of participants in multicast sessions has more possible limitations than the number of participants in broadcast sessions. Therefore, the number of multicast participants varies from a few to possibly all network users.

Communication configurations of broadcast services are not static. Users of broadcast services join and leave sessions freely, like in present TV broadcast services. Characteristic of broadcast services is the wealth of simultaneously offered broadcast channels, which is the case in television services for example. Such a wide range of provided services leads to the well-known phenomenon of 'zapping' users (or 'TV channel swimming') which is characteristic for users searching for the right service. This search process leads to very short session holding times and dynamics in the communication configuration. This is also applicable for multicast services, which can also be used for widespread distribution of television and radio programs.
On the other hand, when the user has found an interesting program, the session holding time can be hours (TV program) or even days or weeks in case of radio channels which are used for background music.

Charging

The possible values of the charging attribute are: subscription, session or sponsor based. Session based charging is not possible for broadcast services, caused by the fact that broadcast services have un unlimited number of unspecified destination points. The alternatives, subscription based and sponsor based charging, introduce an additional requirement: it must be possible to cut off subscribers who refuse to pay or who simply do not want the service to be delivered. This means that such network users must be identified in some way in order to be able to cut them of the broadcast service. An example of a sponsor based broadcast service is a television channel which is financed by the distributor by advertisement yields.

While a broadcast service is delivered to (nearly) all network users in a given network part, multicast services are delivered to a limited number of specified destination points. This means that the number of users can vary from a few to many and that some registration system is required. The fact that the number of destinations is specified for multicast services offers more flexibility in charging methods compared with broadcast services. The possible values of the charging attribute are:

- subscription based
- session based
- sponsor based.

In case of subscription based charging, the subscription contract can be valid during arbitrary time intervals. This offers the possibility to provide a particular TV channel only in the weekend for example. Subscription based services offer also the possibility to implement closed user groups. Closed user groups can be provided by a subscription contract for a selected number of participants. Session based charging requires the account management to administrate for each customer which service is used and how long the service is used. Examples are Pay-TV and Pay-radio.

Communication establishment

In case of broadcast services, any network user is able to participate as long as the service is available. In subscription based multicast services, the consequence for communication establishment is first that communication establishment is only possible during the time intervals agreed in the subscription contract, and second that communication need not be actually established during the subscribed time intervals.
Communication establishment in session based charging services is on demand. Service examples are pay-TV and pay-radio where the customer only pays for the TV or radio channel in use. Another possible session based charging multicast service is Near Video on Demand (NVOD) [CHAN94]. While interactive video on demand services provide a unique video information stream to each viewer with full interactivity and virtual VCR functions, near video on demand services generate multiple streams from the same movie at short, regular intervals of say a few minutes. Viewers gain their VCR functions by jumping between streams. The service can be configured either to allow instantaneous jumping, at the expense of seeing some parts twice, for example, or to save the time stamp of a viewer's interruption and return to it exactly, at the expense of some waiting time. For a video server, interactive video on demand is a more costly service than near video on demand. With less interactivity more viewers can watch a given title simultaneously by taking advantage of multicast capabilities of the network. Depending on the number of viewers, which is not limited, near video on demand could be economical attractive for the service users and for the service provider, especially in the case of popular videos.

2.1.3 Conference services

Conference services have both interactive and distributive characteristics, as was concluded in section 2.1.1. The communication configuration of conference services is multipoint-to-multipoint or point-to-multipoint and the communication symmetry is bidirectional or bidirectional asymmetric. This section investigates the possible values of the remaining service attributes.

User information type

The value of this attribute mainly characterizes the type of service. Possible values are:

- audio/speech (multi-party telephony service)
- files or data (computer group multicast)
- mix (videoconference)

Videoconference and multi-party telephony services are examples of currently provided telecommunication services. Computer group multicast (CGM) is a general term for services using a virtual multimedia space for collaborative work or operations. Examples of computer group multicast services are:

- joint CAD/CAM services
- joint document editing
- electronic mail based chat box
- replicated database updating
- software update distribution

- real time distributed control

Point-multipoint capabilities enable a range of possibilities for remote education services. Traditionally, education sessions are characterized by a group of students mainly listening to the teacher, i.e. the communication is rather unidirectional with sporadic interrupts from the students when they ask questions. Thus the degree of interactivity is low. If remote education is implemented in the similar way as a telecommunication service, then the group of students becomes the group of destination endpoints. Reverse information flows from the destination endpoints to the teacher are required for a limited form of interaction. A major advantage of this form of remote education with limited interaction is the theoretical unlimited student group size. Questions of students can be serviced in sequence on first come basis or with a sort of scanning mechanism giving each student in turn a particular period of time for interaction with the teacher(s).

Except the traditional form of education, also other forms tele-education can be designed. Use of video is not a prerequisite, education sessions can be based on computer generated graphics and text for example. Also mix of video and graphics can be useful for education services.

Communication configuration

For practical reasons, most conference services will have a limited number of participants. For example, videoconferencing and audioconferencing becomes difficult when the number of participants is beyond ten persons for example. Tele-education services possibly have more participants. Depending on the desired degree of interactivity, the number of possible participants varies between a few and a few hundred.

The dynamics of the communication configuration is determined by whether users join or leave existing sessions. This must be possible for all types of conference services. Therefore, the configuration requirements are dynamic. Session holding times can be in the range of seconds for replicated database updating or in the range of hours for videoconferencing.

Communication establishment

Depending on the service type, communication establishment is either semi-permanent in case of subscription based services or 'on demand' when there is no subscription contract. Characteristic for conference services is that communication sessions between all participants start at once, in contrast to distributive service where the participants determine communication establishment in an individual manner. Nevertheless, communication establishment in conference services is also determined by joining or leaving participants.

Quality of service

The interactive character of conference services imposes strict requirements on information transport delay.

Time distribution

The time distribution of several conference services will depend on their penetration and thus economical attractiveness within business and social activities. From the current equivalent of interactive communication, telephony, a sensible expectation would be that videoconference services will be used most frequently during working hours. In first instance, probably videoconferencing will achieve higher penetration in business life than in social life for economical reasons. For computer group multicast the same statements are valid. However, for remote education it is difficult to presume users behaviour patterns.

2.1.4 Overview of requirements

The assignment of values to several service attributes in the previous sections implicates service specific implementation requirements for the transport techniques to be used for the realization of information transporting connections. The diagram in table 2.1 gives an overview of the attributes and their values which have direct consequences for the connection management. The objective of this table is not to give exact specifications, but to indicate main correlations and main dissimilarities between values of several service attributes. The first six service examples in the diagram are unidirectional broadcast and multicast services. The latter four are examples of conference services. The estimated possible values for three attributes are indicated using an interval for the several service examples. The length of the intervals forms a measure of variation.
Considering the number of participants, distributive services have inherently more participants than other services with a limited form of interaction. Broadcast services speak for themselves, differences between numbers of users of subscription services and pay per session services depend on economical attractiveness to the customer and the variety of channels offered.

Regarding the session holding times, characteristic of distributive services is their short holding times in case of 'channel swimming', designated with one asterix in the third column of table 2.1. Furthermore, the session holding times of distributive services will depend on the duration of programs and user behavioral patterns. Therefore, the variation in session holding times will be large.
The column of 'Configuration Dynamics' indicates that most point-multipoint services have (to some degree) dynamic characteristics with respect to user join and leave events. Of major importance is the conclusion that both classes of multipoint services require connections with a *varying* number of users.

Table 2.1 Approximation of correlations and deviations in service characteristics

Service types	Number of participants	Session Holding Time	Configuration Dynamics
	few *many*	*seconds* *hours*	*static* *dynamic*
Broadcast TV	******	* ************	******
Pay TV	**********	* **************	*******
Subscription TV	*************	* **************	******
Broadcast Radio	*************	* *******	********
Pay Radio	********	* **********	********
Subscription Radio	********	* **********	********
NVOD	****	*******	***
Video conferencing	**	*************	***********
Remote education	***********	*************	**************
CGM	***	************************	********

2.2 Multipoint connections

Information flows supporting telecommunication services are realized in the network by set up of connections between communicating users attached to the network. This section characterizes distinct connection types by their main attributes and their values and investigates which values are required to implement multipoint services as described in the previous sections.

2.2.1 Connection attributes

Conform the description of telecommunication services, also connection types supporting these services can be described using the attribute technique. The list of attributes described here is not exhaustive but concentrates on the following main attributes [I140]:

- information transfer mode
- connection establishment
- connection symmetry
- connection configuration
- traffic type
- connection performance

Information transfer mode

This attribute describes the operation mode for transferring (transporting and switching) user and control information through the network. Possible values are:

- circuit mode
- packet mode
- ATM (Asynchronous Transfer Mode)

The circuit transfer mode reserves and allocates network resources for a particular connection for a given time interval. The packet transfer mode transports user and control information in packets from source to destination(s) without reservation and allocation of network resources. Control information for routing is added to the packet. The packet size is not fixed. The Asynchronous Transfer Mode (ATM) is a combination of circuit and packet transfer modes. ATM provides packet based information transport with reservation and allocation of network resources (see chapter 3 for a more extended description of ATM). The circuit transfer mode and ATM are connection oriented transfer modes, the packet mode is a connectionless transfer mode.
Connectionless transfer modes are not considered in the sequel of this document. Therefore, the following description of connection attributes presume connection oriented transfer modes.

Connection establishment

The connection establishment attribute describes the establishment mode used to establish and release a given connection. Possible values are:

- on demand
- semi-permanent
- permanent

On demand establishment of connections is required in case that a connection must be set up as soon as possible after the set up request (e.g. a users service request). In case of semi-permanent connection establishment, connections or connection elements between fixed source and destination points may be provided for an indefinite period of time after subscription, for a fixed period or for agreed periods during a day, week or other interval and with fixed traffic characteristics. In case of permanent establishment, permanent connections or connection elements are available to connected subscribers at any time during the period of subscription between fixed network destination points requested by the subscribers.

Connection symmetry

The symmetry attribute of a connection has the following possible values:

- unidirectional
- bidirectional symmetric
- bidirectional asymmetric

Connection configuration

The connection configuration describes the spatial arrangement for transferring information between two or more access points. Possible values are:

- point-to-point
- point-to-multipoint
- multipoint-to-multipoint

The topology of a connection configuration is determined by the number of connection end-points and their geographical locations. The dynamics of a connection configuration is determined by end-point add and end-point remove events while the connection is in progress.

Traffic type

The traffic type contains the following subattributes:

- peak bit rate
- mean bit rate
- burstiness

Two distinct traffic types are bit streams with constant (CBR) or variable bit rates (VBR). In constant bit streams, the peak bit rate equals the mean bit rate.

Connection performance

This attribute defines the connection performance in terms of information errors, information losses, transport delay of information and variation in transport delay.

2.2.2 Possible multipoint connection types

This section identifies the possible values of connection attributes which meet the requirements of multipoint services.

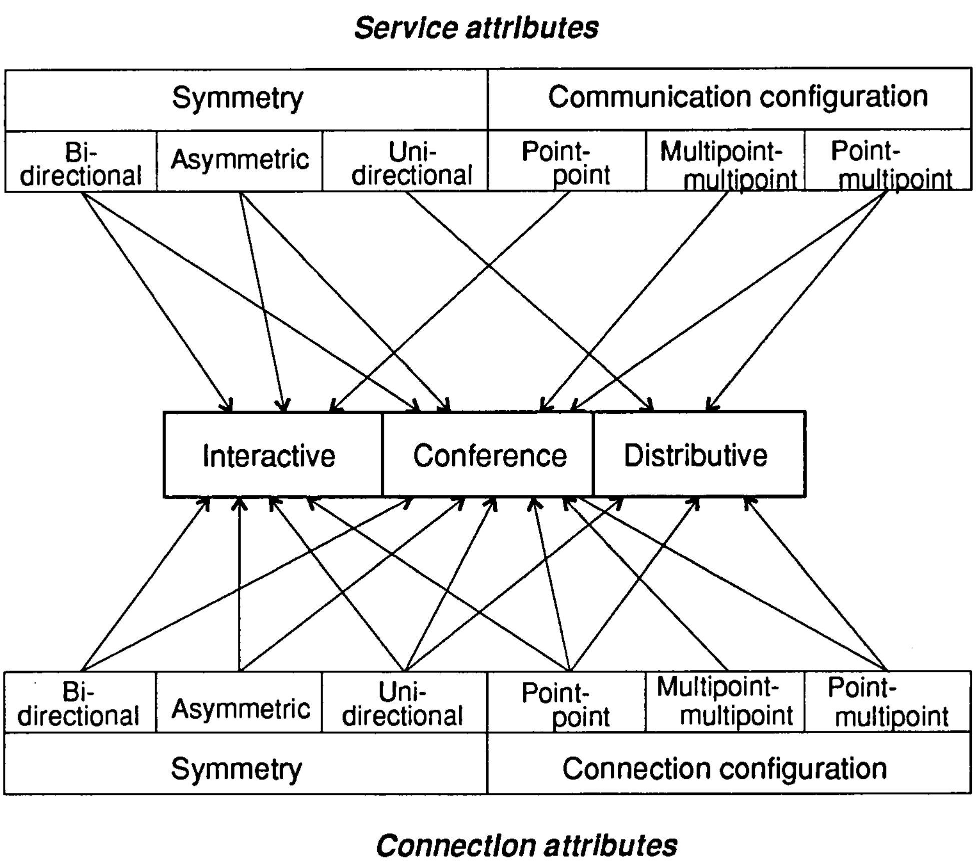


***Figure 4** Connection types supporting the distinct service classes*

Symmetry and configuration

The possible values of the connection symmetry and connection configuration attributes are given in figure 3. Figure 5 illustrates with two distinct connection types for an interactive p-p service that bidirectional communication not necessarily requires bidirectional connections. Figure 6 illustrates the possible connection types for conference services. The optional connection types for multipoint-multipoint communication are:

1. bidirectional asymmetric p-p connections
2. unidirectional p-p connections combined with a unidirectional p-mp connection
3. unidirectional p-mp connections
4. one bidirectional asymmetric mp-mp connection.

A basic requirement for the second and last solution is that the information streams originating from all participants (or from a set of participants) are collected and mixed in one multicast

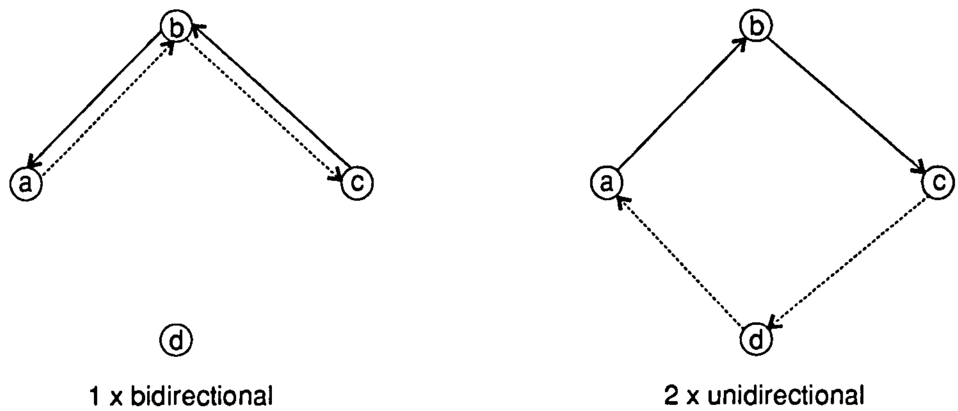


Figure 5 *Examples of bidirectional communication using distinct connection types*

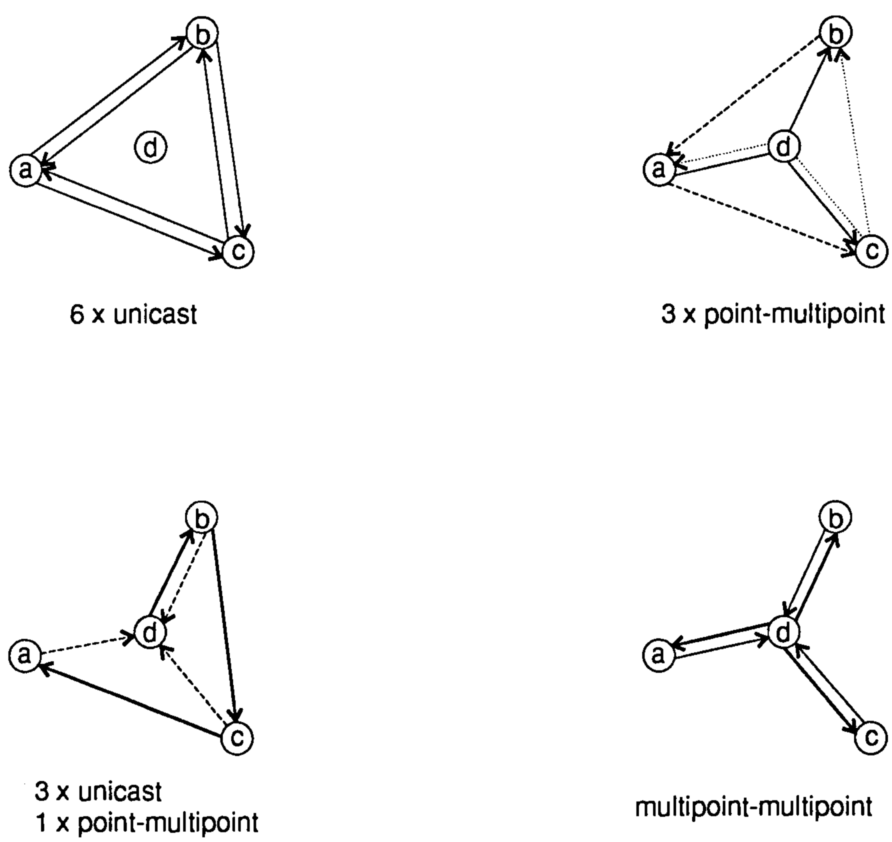


Figure 6 *Distinct connection types supporting an interactive conference service with three participants*

stream. Within the mixed information stream, several channels from different participants must be identified in some way.

The first connection type requires also a collecting and mixing function, but in this case the mixed information stream can be synthesized according to the demands of the distinct participants. In this case, the mixed information stream will not contain information originating from the participant it is synthesized for. The third is the only connection type which does not require a collecting and mixing function.

The possible connection types for broadcast and multicast services are unidirectional point-point and unidirectional point-multipoint connections. For efficiency reasons, p-p connections are of minor importance compared with p-mp connections, see chapter 4 for a further analysis.

Connection establishment

Communication establishment requirements can also be met with several possible values of the connection establishment attribute. It is evident that permanent communication requires a permanent connection. However, communication on demand can be realized with on demand, semi-permanent or permanent connections. An example could be broadcast TV, where the user is connected (the connection is extended) at the moment he selects a specific TV channel. This example illustrates that the connection establishment attribute possibly differs for extension of connections and the set up of complete connections at once.
On demand conference services require on demand establishment of a complete connection to all participants. While a session is in progress, new joining participants require on demand extensions of the existing connections.
Generally, subscription based services do not require on demand establishment of connections because it is known before which connection must be set up.

2.3 Traffic type specific requirements

Transport of information through a telecommunication network requires reservation and allocation of network resources. Service specific traffic types mainly determine how much network resources are required for the connections. Therefore, this section investigates traffic type specific network requirements of video information streams, audio information streams and graphic and text information streams.

2.3.1 Video information streams

Video representations will play a major role in multipoint services, like broadcast and multicast TV or videoconferencing. Therefore, the characteristics of video information streams will be of major importance for network design and connection management. With the video compression techniques currently developed, a range of possibilities arises for transport of video information. The resulting requirements on transport techniques and transport capacity generally depend on the used compression (coding) technique, the required picture quality and the amount of movement in the displayed video scene. Higher picture qualities will demand higher information rates. Complex coding techniques with long processing times will achieve a higher coding efficiency than fast, time critical techniques. Video scenes with low movement degree gain high coding efficiency against scenes with high movement degree (e.g. sports) on which the same coding algorithm will perform poorly.
These dependencies will be investigated in the following, including two appropriate transport modes for the transport of video information. Both modes have significant distinctions in their traffic characteristics insinuating different network design objectives.
Another matter of interest which has also consequences for the capacity requirements on connections, is video service interworking. This term is used for the widely recognized [I211] objective of interworking capabilities between video coding techniques performing on different qualities and distinct formats of video pictures.

Main characteristics of video information streams

Efficient usage of available bandwidth for the transport of video information requires the use of compression techniques which reduce the required information (bit) rate. The currently developed video compression techniques differ in implemented coding algorithm and possibilities to compress video bit streams at distinct levels of picture quality.
The bit stream generated by a coding technique can either be constant or variable, depending on the implemented coding algorithm. Variable bit rate (VBR) coding is a coding method which produces a stream of bits whose rate is time varying according to the variation of amount of information in the original signal (scene content). Thus increasing scene content (more movements) in a given scene will produce an increased bit rate which enables the codec system to maintain a fixed level of picture quality. On the other hand, in the case of constant bit rate (CBR) coding, the produced bit rate must remain constant in spite of increased information content from the original signal, which leads to picture quality degradation.
Possible advantages of VBR coding are:

- Redundant data which may need to be sent in CBR coding is not sent in VBR coding
- Because very little data is transmitted when the information content is low, and because high rates are used only when necessary, VBR codecs are expected to provide an overall higher quality at a lower average rate than CBR codecs.

A disadvantage of VBR coding are the increasing buffer requirements in the transport network to absorb the inherent burstiness of the VBR stream.
The advantage of deterministic CBR information streams lies in the fact that they are easy to manage. A constant amount of network resources are reserved instead of the probabilistic approach needed for VBR. The disadvantage of CBR coding is the required use of increased output buffers in the encoder to eliminate the bit rate fluctuations which introduces extra delay.

Video service interworking

The telecommunication network must be capable of supporting a range of video applications covering a range of resolutions (e.g. from video telephony to High Definition Television (HDTV)). Compatibility between several picture formats plays an important role in the area of video coding techniques. The reason is that higher efficiency can be reached in using network resources if the information stream supporting a given video service can also be used for almost the same video service but at a lower picture quality level (HDTV and standard quality TV for example). Another argument is that only one codec is needed for several levels of quality.
In fact, there are two usable methods to achieve video service interworking:

- Simulcast approach: Transmitting terminals contain multiple encoders, operating at a variety of resolutions and quality levels so that broad inter-connectivity can be achieved by transmitting multiple parallel encoded signals.

- (Hierarchical) Layered signal approach: A layered representation of the video signal is defined. Coders transmit a baseband signal which provides a basic quality service. Incremental signals, which can be used along with the baseband to recover a high quality, are also transmitted. Receiving terminals utilize the baseband and an appropriate number of incremental signals to recover the video signal to the quality which they are capable of displaying.

The simulcast approach does not make efficient use of bandwidth but is simple to implement. The layered signal approach seems suited to a wide range of applications but will lead to higher complexity in codecs.

Coding techniques

A number of video coding techniques have already been standardized or are currently being standardized. The H.261 coding algorithm was developed for use with STM networks [H261], thus it was optimized to operate at fixed bit rates. This coding method uses bit of p * 64 kbit/s, where p is in the range of 1 to 30. The H.261 algorithm covers the use of 525- and 625 line television standards.

An experts group in ISO, known as Moving Picture Experts Group (MPEG) is developing standards for generic applications, initially in particular for storage applications. The first result, the MPEG1 video/audio codec, has become an international standard (ISO/IEC 11172). It has been optimized for constant bit rates between 1.1 Mbit/s and 1.5 Mbit/s. Another system for video and audio coding is currently under development: the MPEG2 video/audio codec. This standard is being optimized for bit rates between 4 Mbit/s and 9 Mbit/s (both VBR and CBR) and will support a range of picture formats including HDTV. MPEG2 HDTV requires bit rates between 10 and 20 Mbit/s [COST93]. The MPEG2 system uses a hierarchial layered signal approach for interworking between distinct picture formats. Appendix D further explains the used coding techniques in MPEG codec systems and the effect on coding and decoding delays.

VBR versus CBR

The MPEG2 coding mechanism produces inherently variable bit rates. However, the coding mechanism is optionally extended with a control mechanism which transforms the variable bit stream into a constant bit stream. Thus both VBR and CBR video information streams can be produced.

Currently, many papers deal with the question whether VBR video coding or CBR video coding should be used in a telecommunication network [KAW93, SIM903]. Maybe the gains of VBR will be too small compared to the additional costs on the network side. An important point is the reaction of different coding schemes on information loss. The multiplexing gain of several VBR channels depends on the tolerable information loss rate.

Although the development of multi-layered coding schemes is still in progress, several papers give strong hints that VBR coding for typical TV sequences will not result in a substantial multiplexing gain [SIM901, -902, -903]. It is stated that perceptive picture quality can significantly differ from the measured Signal to Noise Ratio (SNR), especially in the case of scenes with violent motion (e.g. zooming and scene cuts). Furthermore, it has to be taken into account that if the user of CBR equipment is willing to accept slight quality impairments for let's say 1 or 2% of the time (e.g. in the case of scene cuts), the corresponding multiplexing gain of a VBR source with the same (slightly reduced) quality will be significantly decreased further.

Effect of scene content

An important factor in determining the characteristics of video traffic is the scene content of the input video sequence. The scene-content-related factors, which can vary across video sequences, include colour, brightness and motion content. In [PAN94], the influence of these factors on the actual bit rate were investigated using the MPEG coding algorithm (VBR, restricted to intraframe and predictive interframe coding, see Appendix D). The video input sequences had a frame resolution of 512 x 480 pixels, which is similar to NTSC broadcast quality. The digital output was packetized in cells with a payload of 48 bytes. To illustrate the statistics of distinct video sequences, the results of the experiment are shown in table 2.2.

Table 2.2. *Cells per frame statistics for several sequences*

Sequence name	Cells per frame		
	Peak	Mean	Standard deviation
Indiana Jones 1	931	405	113
Indiana Jones 2	1452	501	204
Last Emperor 1	1095	604	223
Last Emperor 2	809	352	110
League of our own	1252	537	172
Taxi Driver 1	743	242	144
Taxi Driver 2	1001	356	161
Boxing	1337	781	188
News	852	248	113

The corresponding average bit rates vary between 2.1 Mb/s for "News" to 7.8 Mb/s for "Boxing". The activity contained in a sequence is closely related to the average bit rates that are required to encode the sequences.

2.3.2 Non-video information streams

This section investigates traffic specific requirements of information retrieval services and audio services. Information retrieval services can be provided using text character sets or graphical representations. The bandwidth requirements of these services depend on the used representation mode and the amount of information which has to be transmitted. An example of a present information retrieval service is Teletext.

Information retrieval services

Information retrieval services can be provided by using text character sets, graphics and still images, or software-like information types. The bandwidth requirements depend on the used representation mode and the type of service. At this moment, software-like applications are not defined yet. Investigation is needed on this point before a profound prediction on bandwidth requirements can be made.

Although other (Teletext-like) distribution applications are not defined either, it is possible to approximate the bandwidth requirements based on bit rates produced by their form of representation. Probably these distribution services use cyclical transmitted sequences of information entities like in the Teletext service. These information entities will either contain text or image information. For the text mode, the required bit rates can be deduced from the number of characters per page. For example, a Teletext page requires approximately 7 kbits to fill the screen with text (or graphical) characters. Therefore, the required number of bits for pages with text will be in the order of 10 kbits per frame.

In case of using still images, efficient transmission of high resolution images requires compression techniques to lower the produced bit rates. One of the most widely accepted and supported image compression methods is developed by the Joint Photographic Experts Group

(JPEG). The resulting coding algorithm is used for several applications which require digital storage of still images. The results were also employed for the development of the video compression algorithm in the H.261 standard [H216]. Based on the assumption that it would not be efficient to have a decoder for still images and a decoder for video images, a scheme for still image coding has been adopted in the H.261 standard. This scheme was investigated in [LOUI93] using the CCIR 601 frame type with resolution 720 x 486. The resulting average numbers of bits for good quality pictures appeared to be in the range between 6 and 7 kbits per frame.

The conclusion is that the bandwidth requirements of text or graphics based information retrieval services are far below those of video based services.

Audio services

The ISO /IEC 11172 MPEG (layer 2) audio coder/decoder achieves near-CD quality with a resulting bit rate of 128 kbit/s per channel (thus 256 kbit/s for stereo). It was concluded in [CCIR] that the performance of this audio codec is sufficient for audio distribution purposes.

2.4 Conclusions

Telecommunication services can be described by a number of service attributes. Four service attributes are of main importance for the set up of connections in a network. They are the communication configuration, the communication symmetry, the communication establishment and the information transfer rate attributes. The former two attributes determine the classification of services in a class of point-to-point interactive services, a class of multipoint-to-multipoint interactive conference services and a class of point-to-multipoint distributive services.
Interactive conference services are characterized by a limited number of participants, in contrast to distributive services which have an unlimited number of participants in the case of broadcast. The number of participants in multicast services is limited in case of closed user groups, but can also be large in case of subscription based distribution services.
Conference services require bidirectional information streams, which can be realized with several connection types: point-to-multipoint, multipoint-to-multipoint or a combination of point-to-point connections. Distributive services require unidirectional point-to-multipoint connections.
Communication establishment in conference services requires the network to set up complete multipoint connections on demand. Once the connection is in progress, on demand extensions and removals must be possible caused by joining or leaving participants. Communication establishment in distributive services does not require the network to set up a complete point-to-multipoint connection on demand, complete connection establishment is either permanent or semi-permanent. However, distributive services do require on demand connection extensions.
Session holding times vary between a few seconds and hours for both classes of multipoint services. Session holding times can be extremely short in case of zapping TV watchers.
Video services have the highest bandwidth requirements, varying between 1.5 Mbit/s for an acceptable picture quality and 4 Mbit/s for good picture quality. High definition television requires an information stream between 10 and 20 Mbit/s. These information rates are made possible by MPEG video coding techniques, which produce VBR or CBR bit streams. Good

quality audio requires an information rate of 256 kbit/s for two channels (stereo) with MPEG audio coding.

3 ATM routing strategies

This chapter determines how several types of multipoint connections can be implemented in a network according to the requirements deduced in the previous chapter. Distinct routing strategies will be identified, which meet the general objective of optimal utilization of network resources for services with diverse characteristics. From existing telecommunication networks it is kwown that connection routing must adapt to actual network loads in order to prevent early congestion or blocking of specific network parts.

Connection management requirements differ for several network parts. Therefore, first a generic telecommunication network architecture will be given. Using the currently developed transport technique Asynchronous Transfer Mode (ATM), multipoint connection routing is implemented in the Virtual Path and Virtual Channel concepts using the given network architectures. Of main importance are the distinct requirements of Virtual Path and Virtual Channel solutions onto the multipoint connection management.

3.1 Description of telecommunication network architectures

3.1.1 Definition of a telecommunication network

The main function of a telecommunication network is to interconnect two or more remote customers who want to communicate with each other by means of telecommunication services. The network is responsible for reliable transport of information in an economical manner.

A telecommunication network generally consists of an architecture of *nodes* and *connecting components* between those nodes. The communicating customers are attached to these nodes. A network architecture can be described using the following elements [STIE91, BREED92]:

- The *location* of the network nodes
  Locating network nodes is an optimization problem depending among others on available locations and traffic patterns.

- *Interconnection patterns* between the network nodes
  There exists only a limited number of generic interconnection patterns: star, tree, bus, ring and mesh structures.

- The *functionality* of the network nodes
  Examples of network nodes are switches, concentrators, cross-connects, bridges and routers. In fact, the operation of the network is determined by the functionality of the network nodes.

- The *physical medium* constructing the connecting components
  Generally, the choice of the physical medium for the connections, like fibre or coax, has no direct implications on the design of the network architecture. However, the physical medium possibly determines the geographical reach of the network.

The network architecture can be divided into three parts due to geographical span and traffic concentration. The distinguishable network parts are [STIE91, BREE92]: the *Customer Equipment Network*, the *Access Network* and the *Backbone* or *Trunk* network. Figure 7 illustrates these distinct network parts.

- *Customer Equipment Network*
  Customer equipment networks are networks located at domestic and business sites. They provide the transport of information over short distances within one user group, e.g. the employees within one division of a company. Normally these networks fall beyond the scope of the public network.

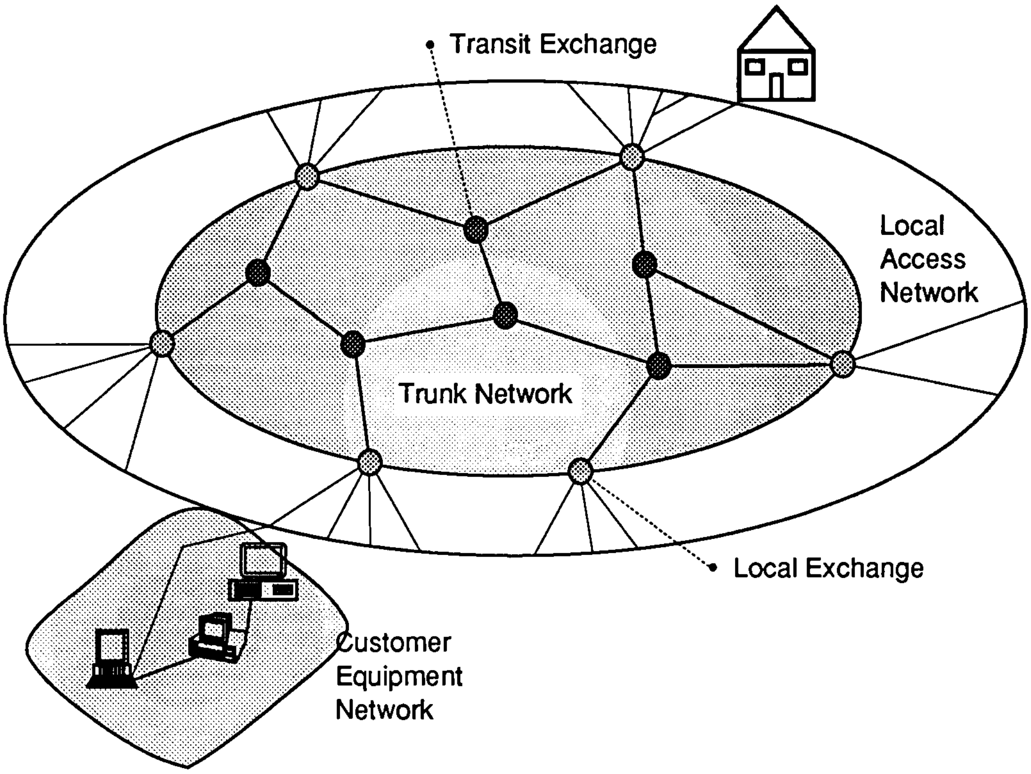
- *Access Network*
  The access network provides customers access to the public network and thus constitutes a part of the public network.
  Generally, the upstream information traffic transmitted by domestic customers and destined for the trunk network is limited. In the opposite direction, distributive services will mainly determine the volume of the downstream traffic, which is delivered by the trunk network and destined for customer equipment networks. In business segments, the traffic in both directions will be small in volume compared with the local traffic intensity and the information streams in the trunk network.

- *Trunk network*
  The trunk network provides the transport of concentrated information streams over long distances, national and international and provides interconnection of the attached access networks.

The interface which separates the trunk network and the access networks, consists of nodes that will be called *Local Exchanges* (LEX). The nodes inside the trunk network not interfacing the access network are the *Transit Exchanges* (TEX), see figure 7. The functionality of both types will be deduced in the following sections.

3.1.2 Physical and functional topologies

The previous section defined a telecommunication network as an architecture of nodes and connecting components between those nodes. This definition applies obviously to physical components which constitute a *physical topology*. However, there is a second level, the *logical* or *functional topology* which is determined by the implemented network techniques independent of the underlying physical topology [STIE91]. This concept of functional topologies independent of the lower level topology might be extended to several levels.
Illustrating for this concept is the use of Wavelength Division Multiplexing (WDM) in fibers. By using distinct wavelengths within one fibre it is possible to construct a functional topology determined by a given wavelength, which is independent of the physical topology.



7 Geographical separation of the network architecture

Interconnection patterns

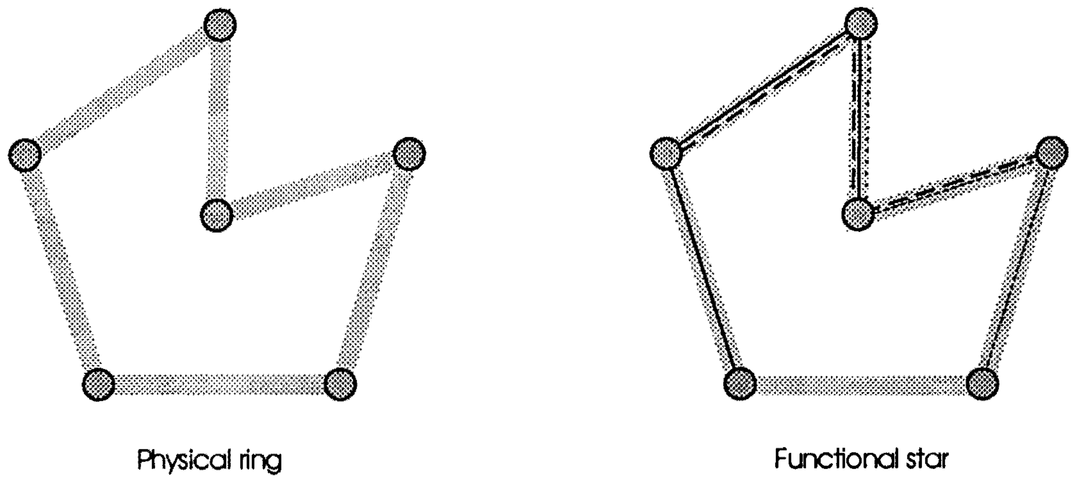
The topology, whether it be at physical or functional level, is determined by its *interconnection pattern* of linking components between the communicating nodes. Interconnection patterns can either be generic or a combination of generic interconnection patterns. With generic interconnection patterns as building blocks, each possible interconnection pattern can be constructed. The generic interconnection patterns are:

- star or tree
- ring
- bus
- meshed

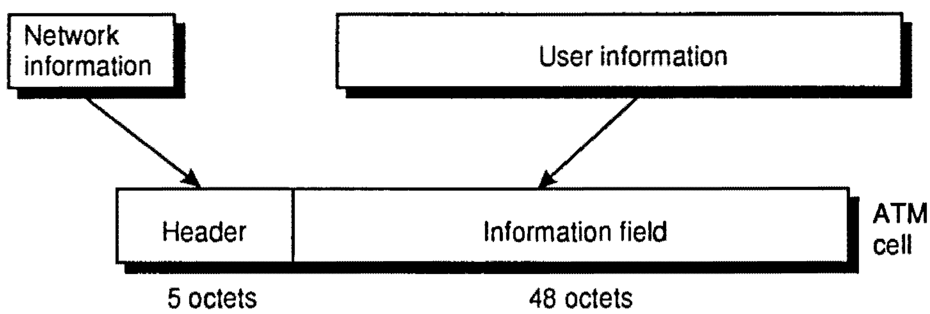
Figure 8 illustrates the separation of physical and functional topologies, using a ring structure at physical level and a star pattern at functional level.

3.2 Description of Asynchronous Transfer Mode

Hitherto, the telecommunication network is described basically in terms of topology and interconnection patterns. However, to offer telecommunication services, some means of transportation of the user information is needed. A transport technique that is currently being developed and which promises to support a wide range of services is the *Asynchronous Transfer Mode* (ATM). The Asynchronous Transfer Mode offers major advantages for the integrated broadband transport of multimedia information and enables realisation of a real multi-service network, including multipoint service capabilities [VRIES94]. The next sections will explore the basic concepts of ATM in order to translate the service requirements to network requirements using ATM.



8 Example of a functional star on a physical ring architecture



9 The ATM cell

3.2.1 The connection concept

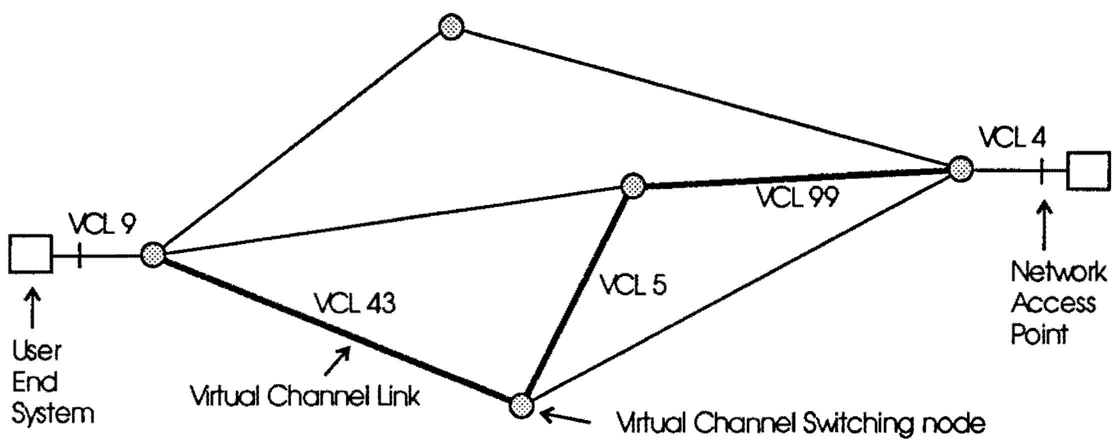
The Asynchronous Transfer Mode uses a unified approach to transport user data of all types of information services. Control, management and user data is assembled into fixed size elementary information units called *cells*. Cells consist of a 5 octet header, which holds network information and a 48 octet information field which is reserved for user information, see figure 9.

ATM is connection oriented, i.e. before data can be transmitted a connection must be set up by negotiations between user and network with respect to the required bandwidth, quality of service etc. At the end of a communication session the connection is released. During the information transfer phase an end-to-end communication path is available. The set up and termination operations require special (network) communication sessions which are called *signalling* or *management* procedures. The signalling procedures use special communication channels preserved for set up and termination procedures.
ATM provides two basic concepts to provide end-to-end connections, namely the *Virtual Channel* concept and the *Virtual Path* concept.

The Virtual Channel (VC) concept

The VC concept provides a way to create unidirectional logical communication links called *Virtual Channel Links* (VCLs) between *VC switches* or ATM network access points. A given

VCL is identified by its *VC Identifier* (VCI), while a VC switch connects VC links and translates the VCI values. The VC Identifier is contained in the headers of the cells. Virtual Channel Identifiers identify unambiguously VCLs and enable the multiplexing of several VCLs on the same physical link.
A Virtual Channel Connection (VCC) is a concatenation of VCLs thereby forming an end-to-end route through the network in order to support the transparent transfer of information between two ATM network access points. A VCC can also have a point-to-multipoint configuration in which there is more than one destination.
Figure 10 illustrates the Virtual Channel concept.



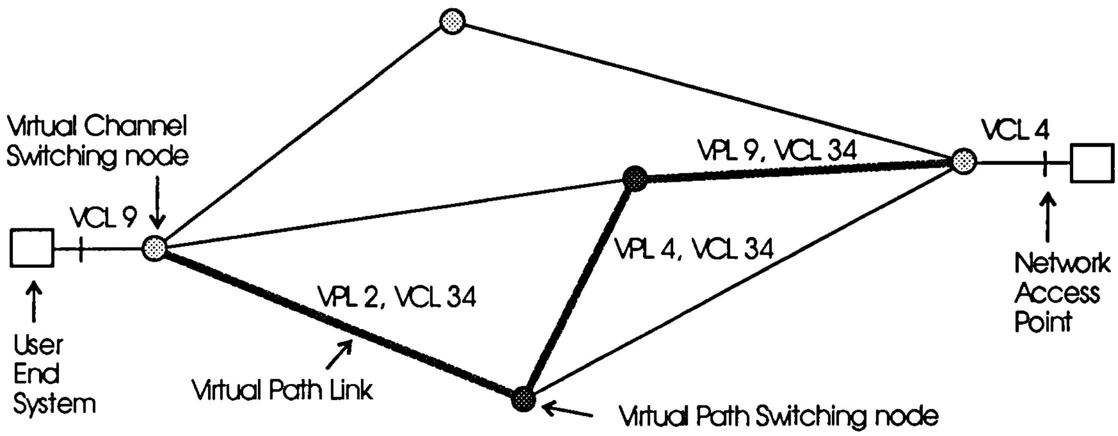
10 The Virtual Channel concept

A VCC is characterized by its bandwidth, quality of service parameters, symmetry and interconnection pattern. The required bandwidth and quality of service parameters mainly depend on the type of traffic (e.g. traffic for video or audio applications), or depend on specific traffic contracts (e.g. data traffic). In ATM all types of traffic are supported, see also Appendix D. The main types are Constant Bit Rate (CBR), Variable Bit Rate (VBR) and Available Bit Rate (ABR). The latter type requires no QoS guarantees (except cell loss) and utilizes the parts of network capacities which are temporary not in use by other types of services requiring a guaranteed QoS.
The symmetry of the VCC is determined by the symmetry of communication, see section 2.1.1. The interconnection pattern depends on the results of routing procedures embodied by the connection management.

The Virtual Path concept

The Virtual Path concept originated after consideration of trends in developing networks. These trends suggest that the control cost in future networks will be a far higher proportion of the overall network cost. The Virtual Path concept helps contain the rising network control cost by grouping connections sharing common paths through the network into a single unit. Network management actions can then be applied to a small number of groups of connections instead of a large number of individual connections. The group of connections is referred to as *virtual path*. The grouping of connections makes the ATM Virtual Path concept resemble cross-connecting in large circuit switched networks (e.g. bundle routing or trunk switching).
The Virtual Path concept introduces *Virtual Path Links* (VPLs) and *Virtual Path Connections* (VPCs). A Virtual Path Link is a logical unidirectional direct link between two neighbouring nodes in a network architecture, which are called *VP cross-connects*. The VP cross-connects connect VP links and translate the *VP Identifiers* contained in the cell headers. The VPC is

again a concatenation of VPLs linked by VP cross-connects. A Virtual Path bundles multiple VCLs or VCCs. Just as a VCC, a VPC also may have a point-to-multipoint configuration. The overall Quality of Service (QoS), i.e. the quality of cell transport in terms of delay, variation in delay, loss, etc., provided by a VPC is the same to all of the VCLs carried within it. Figure 11 illustrates the VP concept.



11 The Virtual Path concept

The VCI relates to call-by-call cell switching functions at VC level (a *call* is defined as a communication session) and enables therefore *real-time* connection set up. The VPI relates to *semi-permanent* cross-connecting functions controlled by network management entities. Figure 12 illustrates the physical use of transmission link transfer capacity in the ATM connection concept. The link transfer capacity is shared by an arbitrary number of VPLs which in turn carry an arbitrary number of VCLs.

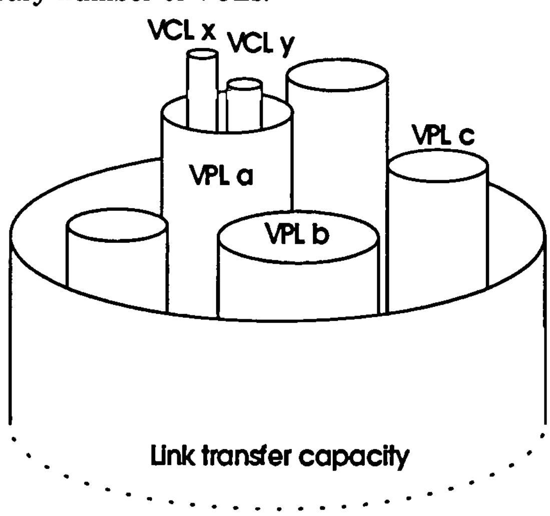


***Figure 12** Dynamic use of transmission link tranfer capacity in ATM*

The Virtual Path concept offers the possibility to create a functional architecture by using the semi-permanent cross-connecting functions *independent* of the physical architecture. To VPLs bandwidth can be assigned independent of the physical capacity. User demands determine much more the assigned bandwidth and not physical constraints.

In contrast to VCCs, the interconnection pattern of VPCs will generally not depend on a single communication session because one VPC contains several VPLs. Thus, routing procedures for single calls are not required at VP level.

3.2.2 Traffic performance aspects

The ATM transport technique introduces some ATM specific aspects which influence the overall Quality of Service (QoS). The QoS is partially determined by the inherent characteristics of cell based transport. ATM cells carrying user information travel through Virtual Channels consisting of several Virtual Channel Links which are concatenated by switching network nodes. These switches collect the incoming cells of different links using cell buffers and translate the channel identifiers in order to transmit cells to the next node (or termination point). Cells passing switching nodes can thereby experience delay, even variation in delay and loss with non-zero probability. Therefore the following measures of traffic performance are defined for ATM networks.

Cell Transfer Delay

A cell transported by the network will experience transfer delay. In ATM a fixed cell transfer delay can not be guaranteed. Due to non-deterministic arrival of cells at cell switching network nodes (VC-switches or VP cross-connects), a random delay component is introduced. This random delay component depends on the loading and traffic mix of the network, the number of switching nodes along the VCC/VPC and the traffic characteristics of the source. This random delay component is called Cell Delay Variation (CDV).

Cell Error Ratio

The non-deterministic arrival of cell at switching nodes necessitates the use of buffers to protect the cells in traffic bursts from being lost. However, when the available buffer space does not suffice due to excessive high traffic peaks or due to a number of simultaneously occurring bursts, cells get lost. The ratio of the number of lost cells during a specific period time to the total number of cells sent during that period, is called the Cell Loss Ratio.

Cell Misinsertion Rate

Due to undetected bit errors that affect the routing information in the cell headers and malfunctioning network nodes, misinserted cells may arrive at the destination. The ATM performance measure Cell Misinsertion Rate can be defined as the total number of misinserted cells observed during a specific time period divided by the time period duration.

3.2.3 Traffic control mechanisms

A number of functions and mechanisms help in managing the traffic offered to the network and providing the customers the QoS they ask for. The most important functions are described below.

Connection Admission Control

Each time a new VCC or VPC is initiated it has to be examined if there are sufficient resources available to accept the new VCC/VPC. The function that performs this examination is called Connection Admission Control (CAC). Connection Admission Control provides a means to preserve the agreed QoS of the already existing connections in the network.

Policing

Policing is the monitoring of traffic in order to determine whether user cell streams behave as has been agreed upon at call set up, and the invocation of specific actions in case of violations or user misbehaviour. The policing function can be performed at two distinguishable interfaces in the network:

- the *User Network Interface* (UNI), this is the interface between terminal equipment of the subscriber and a network termination (network access point)
- the *Network Node Interface* (NNI), this is the interface between two neighbouring network nodes.

If the policing function is performed at the UNI, then it is called Usage Parameter Control (UPC). If performed within the network at NNI, the function is called Network Parameter Control (NPC).

Priority

One bit in the ATM cell header will be used for explicit Cell Loss Priority (CLP) indication. When this bit is set a cell has low priority with regard to loss, i.e. such a cell is subject to discarding depending on network overload and congestion conditions. The ATM service user can select which cells have the higher loss sensitivity.

3.3 Connection management in ATM

Connection management has the task to set up, maintain and release connections in the network. This section identifies the main objectives of connection management which form the basic requirements for the design of routing strategies. Maintenance of connections requires fault management functionality and performance management functionality, which is the subject of section 3.3.2.

3.3.1 Main objectives

A main objective of connection management in general is the efficient use of network resources such that revenues from services offered by the network will be maximized. In other words, connection management contributes to network resource management. In practice, this means that lowering of blocking probabilities due to network overload becomes an integral part of connection management. The choice of connection configurations between communicating points in the network will strongly influence the maximal throughput of the network. One essential function of connection management that determines these configurations is *routing*. Routing procedures select a combination of nodes and links between those nodes constructing the connection which carries the information streams. Thus the *routing algorithms* will mainly determine the maximal throughput of the network. In ATM networks, routing functionality is required for both the VP and the VC level.

Traditionally, public telecommunication networks carried mainly bidirectional (telephony) or mainly unidirectional traffic (CATV). The consequence of an integrated approach is that the network must be capable of transporting a mix of both unidirectional and (asymmetric) bidirectional information streams. As a result, the occupation of transport capacity in forward and backward directions of connecting links between network nodes is not balanced automatically like in telephony networks. Unbalanced occupation of forward and backward link

directions lead to unnecessary blocking of requests for (asymmetric) bidirectional connections. Thus, for efficiency reasons link occupation must be balanced by some mechanism.
The larger the number of connections allocated to a particular link, the easier the balancing mechanisms operate (e.g. in the trunk network where large traffic streams are concentrated). On the other hand, a limited number of connections allocated on a physical link will for that very reason intricate the balancing mechanism. In the case of domination of distributive traffic in the outermost edges of access networks for example, balancing is even not sensible any more. In such cases the physical network architecture must suit the local traffic patterns.

Another main objective of connection management in general is to maximize reliability of information transport through the network. Reliability design objectives can be arranged according to their applicability to distinct network layers. Traditionally, reliability design objectives emphasized network equipment availability. For example, a frequently used measure of reliability at the physical layer is the failure rate of network elements (switching and transmission equipment). With ATM, functional networks are provided by Virtual Paths introducing reliability objectives at the functional level. Due to transmission bit errors and information processing errors, the reliability at the functional level decreases independent of the measure of reliability at the physical level.

As concluded in the previous chapter, distinct services put distinct requirements onto connection management with respect to the number of participants, channel set up speed, adaptability to varying communication configurations and session holding times. However, in ATM networks these requirements will also depend on the implementation in the VP/VC concept. Both the VP and the VC level require connection management functions which will partly overlap but will also partly differ. Implementation choices for the distinct services with VPs or VCs will clear these similarities and differences. Starting in section 3.4, further investigation of implementation possibilities in ATM for multipoint services will clear the specific requirements of distinct service types on VCs and VPs.

3.3.2 Fault management and performance management

For VPCs and VCCs two additional types of management functions have been identified, *fault management* functions and *performance management* functions [I610]. Fault management functions are needed to detect and signal malfunctioning of (parts of) connections. Performance management functions are required to detect degraded performance of VPCs or VCCs, in terms of loss of cells, mutilation of cells, excessive delay of cells, etc. The fault management functions and performance management functions are called *Operation and Maintenance* (OAM) functions.
The OAM functions are to be performed independently at separated levels: the physical level, the ATM VP level and the ATM VC level. In the ITU-reccommendation I610, the OAM flows which carry OAM information are defined bidirectional for each level. The two ATM level OAM flows are referred to respectively by F4 (VP) flow and F5 (VC) flow. Both flows are initiated at or after VCC/VPC set-up.
F4 cells are identified by a fixed VCI value, F5 cells by the Payload Type Identifier (PTI) in the header (see Appendix C). Cells belonging to the F4 flow of a given VPC will have the same VPI as the user cells which travel the VPC. For both directions the same pre-assigned VCI value will be used.

Both OAM F4 and F5 flows may be associated either with the overall end-to-end connection (VPC or VCC) or a part of the connection, e.g. one or multiple inter-connected VPLs/VCLs. OAM flows in the latter case are indicated with *VPC/VCC segment* OAM flows. After connection set up, F4 and F5 cells can be inserted, monitored and extracted at any terminating point of the VPLs respectively VCLs being part of the overall connection. Appendix B further describes specific OAM functions and related problems in point-multipoint connections.

Relationship between fault management and connection management

Connection set up normally takes place after requests from customers. Thus, routing procedures are carried out according to customer demands for connections. However, in the case of malfunctioning of a part of a point-multipoint connection with many users, it is not desirable to clean up the complete connection and then reconstruct a new one because many users will then be affected by a disruption while the fault may only be local. To overcome this problem, the connection must be partly 'repaired'. To repair a malfunctioning existing connection, the faulted part needs to be rerouted while the rest of the connection remains unchanged. This requires special *rerouting features* in the routing strategies of the configuration management.

3.4 Routing

3.4.1 Routing strategies

The previous section explained the essence of proper routing functionality for efficient connection management in general. Also the importance of a balancing mechanism was made clear. This balancing mechanism must be performed by the routing procedures which select a combination of links and nodes in the network architecture constructing p-p or p-mp connections. Evidently, the route selection procedures need network status information giving a measure of balance in link occupancy. Thus, this is the first profound reason to implement network status dependency in routing strategies for asymmetrical connections.

A second basic requirement instigating *state-dependent* routing comes from the main objective of improving network resource utilization. State-dependent routing anticipates on momentary network utilization patterns, improving network throughput [SIB93, ASH90]. State-dependent routing contributes to avoidance of network *congestion* and lowers *blocking* probabilities.

Congestion

The term congestion is generally used to indicate network instabilities or network overload situations. Conform ITU recommendation I.371 congestion is defined here as a state of network elements (e.g. switches, cross-connects and transmission links) in which the network is not able to meet the negotiated network performance objectives for the already established connections and/or for the new connection requests.

In general, congestion can be caused by

- unpredictable statistical fluctuations of traffic flow
- fault conditions within the network (failures, improper functioning admission control mechanisms)

Congestion caused by statistical fluctuation has short term characteristics in contrast to congestion caused by fault conditions which has consequences for longer periods of time.

In addition, momentary congestion caused by statistical fluctuation has a local nature while the impact of fault conditions will generally not remain restricted to specific network parts.
Long term circumstances are related to long term, mean utilization determined by the number of connections admitted per link. Therefore, prevention of long term congestion is a task of admission control mechanisms. State-dependent routing strategies anticipate on decisions of these admission control mechanisms.

Blocking

Blocking of newly requested connections is a result of congestion, whether it is long term or short term congestion. Blocking is the action undertaken by the admission mechanism to protect already established connections from performance deteriorations.

State dependent routing

The importance of state-dependent routing strategies in improving network performance under extreme conditions such as congestion is well-established by experiences in telephony networks, see [ASH90, ASH92, REG90, KEY90]. Routing strategies can anticipate on two forms of network status dependency: short term and long term network status dependency. The following definitions will be used in this report.

- *State-dependent* routing strategies adapt to momentary network status, i.e. short term adaptability to network element occupancy.
- *State-independent* routing strategies disregard momentary network utilization.

In contrast to short term adaptability, routing strategies can also utilize long term network status information. Daily traffic patterns form elementary long term network status input.

- *Time-dependent* routing strategies adapt to chronical (periodical, recurrent) traffic patterns.
- *Time-independent* routing strategies disregard chronical traffic patterns.

State-dependent routing exhibits automatically time-dependent routing. State-independent routing is either time-dependent or time-independent.
Primary paths are the routings selected by state-independent routing strategies whether it be time-dependent or time-independent. A criterium for primary paths for symmetric services could be the minimum-hop criterium for example. The primary path is then simply the shortest path by counting the number of switching nodes included in the route.
Alternate paths are non-primary paths. They are selected by state-dependent routing procedures based on specific utilization degrees of network elements and the amount of asymmetry in utilization of network elements in forward and backward directions. *Alternate routing* is synonym for state-dependent routing.
When there is insufficient capacity on a call's primary path, allowing a call to complete on an alternate path prevents the call from being lost. State-dependent routing may thus be thought of as a scheme that exploits idle resources elsewhere in the network, caused by statistical fluctuations or imbalances of link loads. Thus, state-dependent routing shares resources more freely than state-independent routing [SIB94].

Mutative routing

A main distinction between point-to-point connections for interactive services and multipoint connections for distributive and conference services results from the dynamics in multipoint communication patterns. As concluded in the previous chapter, multipoint services have dynamic user groups caused by user join and user leave events. In order to avoid confusion with the vocabulary for state-dependent and time-dependent routing, adaptation to dynamic user groups will be called *mutative* routing.

> *Mutative* routing adapts to dynamics in user groups of p-mp services.

Routing algorithms
Another main distinction between point-to-point routing and multipoint routing exists in the required routing algorithms which select the most 'attractive' routes from a number of optional combinations of links and nodes. The exact definition of 'attractive' depends on the situation which can be modeled using a particular cost function with an arbitrary number of parameters. Examples of possible parameters are the required bandwidth in forward and backward directions, QoS requirements or priority indicators. Although multipoint connections can also be routed with point-to-point algorithms by constructing a 'bundle' of p-p connections, this is not the most attractive approach with respect to the utilization of network resources (link capacity, buffer capacity, switch control functionality), see [VERM93, BHA83, AMMA93, WAX93, KOMP92]. In chapter 4 algorithms for multipoint routing will be investigated, classified, judged and extended if necessary.

3.4.2 Relationship between routing strategies and service characteristics

A founded choice of routing strategies should be based on particular service characteristics, such as mean session holding times, required transport capacity, permitted delay times and whether the connection establishment is semi-permanent (subscription based) or on demand. These characteristics determine the value of distinct routing strategies in practice. An additional point for consideration form the network link load imbalancies. This paragraph identifies generally holding relationships between distinct routing strategies and several characterizing service parameters.

Session holding time
Generally, the value of state-dependent routing depends on the duration of the session holding times. The motivation is as follows. For short holding times it is intuitively obvious that state-dependent routing has advantages. When there is insufficient bandwidth on a call's primary path, the call is prevented from being lost by using an alternate path. However, alternate paths occupy more network resources than primary paths do. For services with short holding times, it is important that the call will be 'saved', although that means less efficient use of network resources (to some extend). Evidently, calls with long holding times should not be routed less efficiently on alternate paths due to momental peaks in network load caused by statistical fluctuation. Therefore, calls with long holding times should not be routed state-dependent, but time-dependent. Calls with long holding times should also be routed more carefully for the very reason of their long holding times.

The conclusion is that large variation in service holding times pleads for two separated approaches for routing:

- Stress on state-dependent routing for services with short session holding times

- Stress on state-independent, time-dependent routing for services with long session holding times.

Consequentially, separation in two classes requires some criteria to assign calls to one of both. A realistic definition of that criteria will be dictated by network load statistics and is beyond the scope of this report. In any case, a flexible approach is always favourable. A gradual transition between both extremes of state-dependent and state-independent providing a 'fine tune' mechanism will offer more flexibility, e.g. by limiting the amount of alternate paths.

Capacity requirements

In addition to holding times of services, also the required transport capacity (link bandwidth, buffer capacity) of connections should influence the choice of connection configurations [GUP94]. Obviously, connections with high capacity requirements should be routed more carefully with respect to occupation of network resources than connections with relative low requirements. As an example can be given the bandwidth requirement of a telephony call compared to the bandwidth requirement of a video conference call. Thus, required transport capacity can be an additional criterion to apply state-dependent or state-independent routing strategies to given service types.

Connection establishment

Concerning the value of the connection establishment attribute, the following considerations should be made in order to determine if a distinction between on demand and subscription based services is sensible. Subscription based service session should be routed state-independent because:

- connections for subscription based services should not be blocked irrespective the network load
- because connections for subscription based services can efficiently be established using 'sleeping VC's', thus a (time-dependent) routing and connection set up procedure is needed only once.

An argument against state-independent routing of subscription based service is that state-dependent routing of short lived subscription based connections contributes to efficient network utilization, just like 'on demand' connections.

Delay times

Delay times are also of importance for selecting appropriate routing strategies, especially in the case of non-reserved bandwidth services. Available Bit Rate (ABR) services belong to this non-reserved bandwidth service type [VRIES94]. Available Bit Rate services utilize temporary available network resources without QoS guarantees (except cell loss guarantees). Applications of the ABR service type are file transfer or e-mail for example. The lack of a limitation on the permitted delay of information transport is a reason to select a state-independent routing strategy because primary paths occupy less network resources than alternate paths.

The requirements on state-dependent and time-dependent routing strategies for several service characteristics are summarized in table 3.1.

Table 3.1 Overview of routing requirements

	Session Holding Time		Capacity		Establishment		Delay time	
	short	long	low	high	demand	subsc.	short	mod-est
State dependent	*		*		*	*	*	
Time dependent		*		*		*		*

3.5 *Broadcast services in ATM*

This section investigates first how the VP/VC concept can be used to implement distributive services with large numbers of users in ATM. Then the consequences of the implementation for the connection management will be deduced.

3.5.1 Point-to-multipoint VPs for distributive services with many users

Characteristic of broadcast services is their large number of users. Also multicast services possibly have such large number of users. Therefore, broadcast and multicast services with many users will be called here *broadcast-like* services. As a consequence of this large number of users, the probability that a particular Local Exchange which is within the geographical reach of the service is not involved in a distribution session equals or reaches zero. This is characteristic for all broadcast-like services where the probability of at least one involved customer per LEX (which connects a large number of customers to the trunk network) equals or reaches one. The conclusion is that each Virtual Channel used for the transport of broadcast-like service information has the following basic requirement: it must attach all Local Exchanges within the geographical reach of the distribution service.

ATM provides two possible levels for the implementation of p-mp connections, namely the Virtual Path and the Virtual Channel level. A Virtual Channel established in a p-mp Virtual Path has automatically the same p-mp configuration of that Virtual Path. When no Virtual Path p-mp configurations are available, the desired p-mp configuration must be arranged at Virtual Channel level.

There are three main characteristics of VC connections for broadcast-like services which enable the implementation of p-mp Virtual Paths:

- the individual p-mp VCs have identical destination endpoints, i.e. the Local Exchanges
- the types of traffic carried by p-mp VCs for distributive services can be classified in a few classes, each class is a collection of VCs with common information transfer rate requirements and common QoS requirements, i.e. a video class for TV and an audio class for radio services (application of layered video coding and decoding systems such as MPEG2 provides the possibility to integrate distinct video picture qualities within one distribution channel, see section 2.3)

• the semi-permanent character of distributive services matches the semi-permanent character of VP management and control functions.

In case of broadcast-like services, the p-mp Virtual Path approach has a number of advantages above the p-mp Virtual Channel approach. The advantages of using VPs for the bundling of a number of VCs were already described in section 3.2. The main advantage of the use of VPs is that management functions can be applied to a number of VCs simultaneously while improving management efficiency in general (e.g. routing, performance management and VC connection establishment).

Another advantage of the p-mp VP approach for broadcast-like services is that (semi)-permanent traffic for distributive services in the trunk network is separated from other types of traffic for interactive, 'on demand' services which cause fluctuations in network load patterns. These fluctuations in traffic intensity enhance network congestion probabilities. A separated approach for distributive traffic should be favoured because large numbers of users will be affected if distributive connections suffer from network congestion while the number of affected users in case of interactive service is limited. It can be concluded that the p-mp VP approach improves reliability of information transport for distributive services.

Figure 13 illustrates the implementation of broadcast-like services in ATM using the p-mp VP approach with an example of two p-mp VP's.

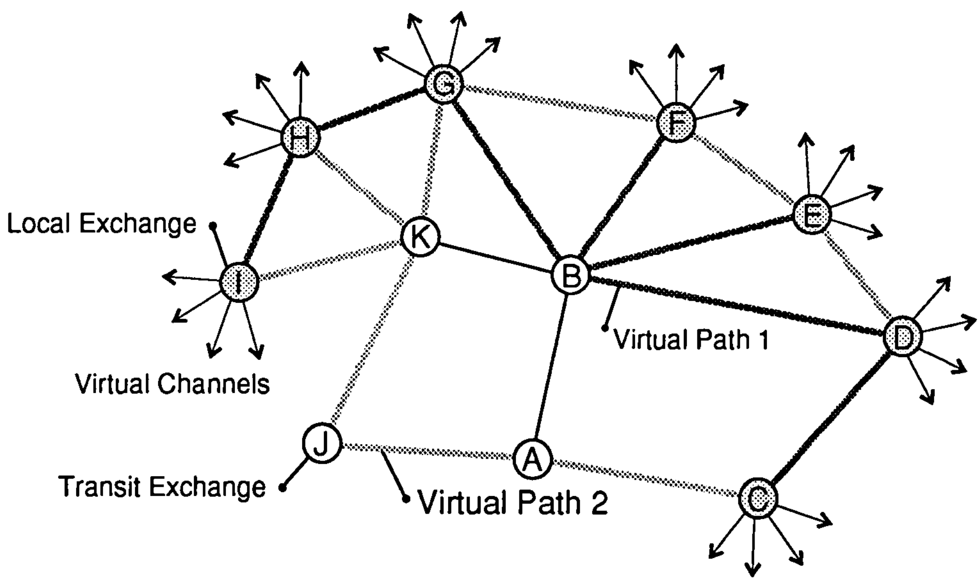


***Figure 13** Virtual Path LEX broadcast configuration*

3.5.2 Requirements onto multipoint Virtual Path connection management

This subsection deduces the consequences of the p-mp VP approach for the connection management and describes how efficiency and reliability for broadcast-like services can be improved.

The (semi-)permanent character of p-mp VPs for broadcast-like services makes the use of time-dependent routing strategies at VP level imperative. Another direct consequence of this approach is that at VC level within the trunk network no routing functionality is required and that p-mp VC establishment is simplified. Connection extensions and removals required by user

join and leave events have only implications for the access network. Thus, no mutative routing is required in the trunk network.
The source endpoint (root) of a p-mp VP is not necessarily the source of a particular p-mp VC carried by that p-mp VP. In addition, a specific p-mp VC can be supported by several p-mp VPs with distinct geographical reach.

Number of VCs within one p-mp VP

Concerning the number of VCs within one p-mp VP, a larger number of VCs improves efficiency in two ways:

- management and control efficiency is improved with grouping of VCs
- in case of Variable Bit Rate information streams, the number of VCs determines the multiplexing gain, see section 2.3.

Number of p-mp VPs

The number of implemented p-mp VPs should depend on the following issues:

- desired symmetry in link loads (especially important when unidirectional distributive traffic dominates)
- the geographical reach of one p-mp VP
- in case that distinct traffic types are separated in distinct VPs, the number of QoS classes supporting distinct traffic types
- desired path diversity for reliability objectives [TAKA94]
- diversity in locations of sources of distributive traffic

Virtual Path intermediate exit

The p-mp VP bundles the VC connections carrying the user information for a particular multicast or broadcast service. These VCs become available at the terminating endpoints of the p-mp VP, thus at the Local Exchanges interfacing the access network. A basic requirement on the LEX is that a VP must be terminated and through connected to a possible neighbouring LEX at the same time, see figure 14.

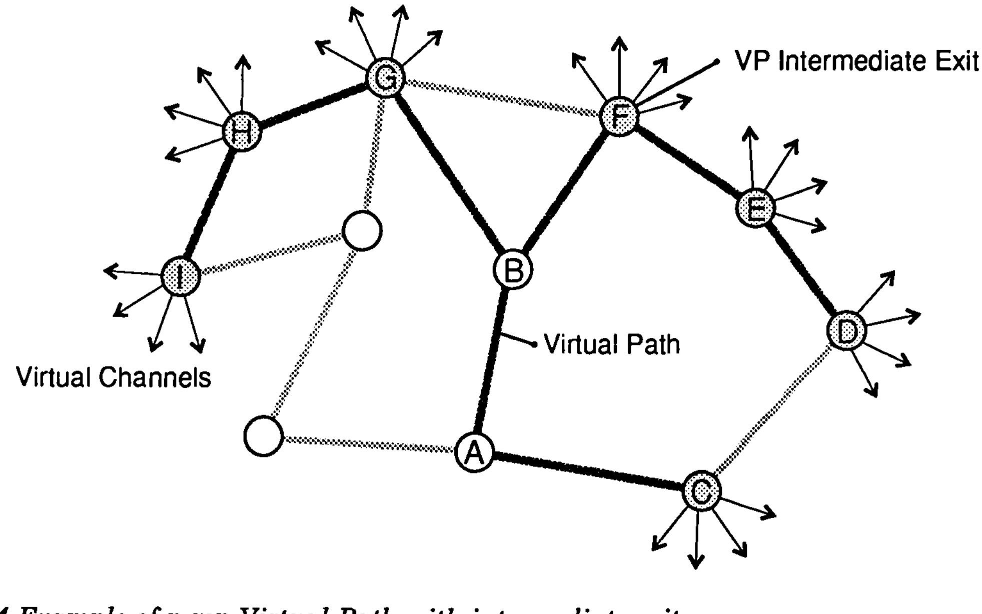


Figure 14 Example of p-mp Virtual Path with intermediate exits

3.6 Broadcast services in access networks

The previous sections mainly stressed on multipoint ATM connection management in the trunk network. This section investigates ATM connection management in the Access Network for distributive services.
Characteristic of broadcast-like services is that a typical customer will receive a wealth of television channels and radio channels while the number of actually used channels will vary between zero and a few. Therefore, permanent establishment of these distributive connections in the access network does not endorse the general objective of efficient network resource utilization. The following three methods for reducing distributive traffic in access networks should be considered. The first two methods are specific for ATM networks, the third method is also applicable to other (existing) network types.

Restrict allocated bandwidth

This solution comprises normal (semi-)permanent establishment of distributive VC connections from LEX to customer equipment as agreed on subscription contracts. However, leaf parts of p-mp connections not in use are restricted in transmission cell rate by extended use of Network Parameter Control functions. More specific, the mean transmission rate parameter of the service contract must be renegotiated when a customer selects a new distribution channel.
The restriction of transmission rate can either result in no transmission or in a transmission rate guaranteeing some basic QoS sufficient to recognize particular programs. This last option could possibly fasten the channel selection process, for example by combining these channels at 'basic rate' into a mosaic channel. As explained in section 2.3, MPEG-2 video coding techniques support layered coding of video signals. When using MPEG-2 compression techniques, the bit rate of basic layer coded video signals could be employed as the minimum transmission rate

Establish and release

This solution only establishes VC connections when they are actually in use by the customers. Obviously, the establishment and release of connections when customers are 'swimming' through their available wealth of TV channels burdens connection management. This is especially the case when programs become interrupted by advertisement messages or simply when popular programs end.

Use local video/audio server

Extend the access network with a *video/audio server* and use p-p connections. In this scenario customers are connected and serviced through p-p connections by a local video/audio server (entertainment centre, distribution centre). Thus, the video/audio server generates video or audio traffic according to the individual wishes of each particular customer. The individual traffic streams are either generated locally in the case of Video on Demand services or collected from the offered distributive services by the (trunk) network.
This concept of integrating Video on Demand services with distributive services in the access network can be supported by any type of network (ATM, ADSL, CATV) as long as p-p connections for customers can be guaranteed providing a minimum downstream capacity per connection of say 1.5 Mbit/s for MPEG1 video bit streams.

3.7 Conference and multicast services in ATM

While the previous section developed ATM solutions for broadcast-like services, this section will investigate how multipoint services with a limited number of participants should be supported in ATM. Connection management requirements differ significantly from those for broadcast-like services, as will be deduced in the following.

3.7.1 Virtual Channels for multipoint services with a limited number of users

Broadcast-like services, as defined in the previous section, cover the range of distributive services with large numbers of users, such that (nearly) each LEX within the geographical reach is involved in a broadcast-like service session. Unidirectional multicast services with a strongly limited number of users cannot be supported efficiently by p-mp VPs for broadcast-like services. Examples are closed user groups services such as multicast of symposia of scientific conferences, Church TV or foreign TV channels watched by a only few emigrants. They will be called *limited user group* (LUG) multicast services in the sequel of this document. It is evident that such multicast services cannot be supported by broadcast connections.
In addition to LUG multicast services, also interactive conference services are characterized by a limited number of participants and cannot be supported by p-mp VPs. Therefore, both the conference services and LUG multicast services require another approach.

Connection types

Whereas distributive multicast services require unidirectional p-mp VC connections, interactive conference services can be supported by distinct connection types (see section 2.2). They are:

- combination of bidirectional p-p connections
- combination of one (downstream) p-mp connection and (upstream) unidirectional p-p connections
- combination of unidirectional p-mp connections, one for each participant
- one bidirectional mp-mp connection.

The following advantages and disadvantages should be considered when a selection of one these solutions is made.
The first option is simple but takes no advantage of multipoint capabilities which offer higher efficiency in network resource utilization. The second solution has especially advantages when the level of interaction between participants is low, like in tele-education services with relative large numbers of participants who ask now and then questions. The upstream p-p connections can be established independently of the permanent downstream p-mp connection when necessary.
The third solution becomes inefficient when the number of participants increases caused by increasing signalling overhead, traffic control requirements etc. The latter solution combines upstream and downstream information flows within one bidirectional connection. The disadvantage is that network resource requirements are not fixed in the upstream direction caused by concentration of information flows transmitted by distinct participants as they reach the root of the connection. This effect causes higher requirements on resource allocation mechanisms and on traffic control mechanisms especially in case of VBR information streams.

It can be concluded here that the choice of one the optional solutions should be based on the following service characteristics:

- the number of participants

- level of interaction between participants
- variation in capacity requirements of upstream information flows

Point-multipoint Virtual Channels
The connection types described above can be supported with p-mp VCs independently of connection configurations at VP level. The VC concept offers flexibility in support of distinct traffic types, distinct QoS levels and on demand or semi-permanent connection establishment. Therefore, p-mp VCs should be used for conference services and LUG multicast services.

3.7.2 Requirements onto multipoint Virtual Channel connection management

This subsection describes the requirements of conference services and limited multicast services on connection management.
In contrast to VP routing for broadcast-like services, routing of connections applies to the VC level in this case. Whereas broadcast-like services have mostly a (semi-)permanent character, the class of conference services contains services with both 'on demand' and (semi)-permanent requirements with varying session holding times and varying transport capacity requirements. Based on the conclusions of section 3.4, these different service characteristics require either state-dependent or time-dependent routing strategies. Thus, both strategies are required at VC level in order to support conference services and LUG multicast services according to the general network objectives.

Connection extensions and partly removals supporting participant join or leave events have implications for the entire connection and are thus not limited to the access network, as was the case for broadcast-like services. The conclusion is that a mutative routing strategy is required for conference and LUG multicast services.
Generally, asymmetrical connections for conference and LUG multicast services will load network links also asymmetrical. Therefore, some load balancing mechanism must be integrated in the routing procedures for these services.

3.8 Conclusions

Connection management plays an important role for network resource management in general. Basic objectives of network resource management in general are high utilization of network resources and low network congestion probabilities and at the same time maintenance of Quality of Service guarantees. Set up of connections requires network resources to be allocated along the route computed by routing procedures. In order to prevent network congestion, routing procedures should adapt to fluctuations in network load.
State-dependent routing strategies are required to adapt the routing of connections to traffic fluctuations at short term. Time-dependent routing strategies are required to adapt routing procedures to long term, periodical traffic patterns. Connections should be routed state-dependent or time-dependent, depending on session holding times, transport capacity requirements, mode of establishment (on demand or subscription) and permitted transport delay times. Characteristic of multipoint services is their asymmetric communication pattern

which causes unbalanced link loads, unless a balancing mechanism is integrated in the routing procedures. Dynamics in groups of service participants require mutative multipoint routing.
ATM supports all types of services including multipoint services and provides fault and performance capabilities at the ATM layer. Broadcast-like services are efficiently supported by p-mp Virtual Paths in the trunk network, which connect all local exchanges within the geographical reach of a distribution service. The (semi-)permanent character of these services requires a time-dependent routing strategy for p-mp Virtual Path connections. User join and leave events have only implications for the access network.
Characteristic for limited user group multicast and conference services is the limited number of participants. They require a distinct approach for the set up of p-mp Virtual Channel connections. Multicast services require unidirectional p-mp connections while conference service can be supported by several connection types. The choice of a connection type should depend on the number of participants, the level of interactivity and service specific traffic variations during a session. Due to the variety in service characteristics of conference and multicast services, both the time-dependent and state-dependent routing strategies are required for efficient routing of their connections. The user group dynamics necessitates mutative routing which possibly affects the entire multipoint connection.

4 Multiple destination routing

A main function of connection management in general is the determination of proper routes through the network. The complexity of point-multipoint routing problems equals the complexity of the well known 'travelling salesman problem'. This means that the computation time for an optimal solution grows exponentially with an increasing number of destinations. Therefore, point-multipoint routing heuristics are required which generate near-optimal solutions in a bounded time. These algorithms must create connection patterns which meet particular service requirements, like 'adding' and 'leaving', and which can be implemented within either state-dependent or time-dependent routing strategies in order to fulfil other previously described objectives.

4.1 Multiple destination connection management

4.1.1 Introduction

In general, a telecommunication network provides a range of combinations of links and nodes providing ways to construct connections between network users. The determination of proper routes through the telecommunication network is a main control function performed by the connection management whether it concerns p-p connections, p-mp connections or mp-mp connections. Services with distinct characteristics require distinct criterias in determining 'proper' routes, as concluded in the previous chapter. These criterias depend on the routing strategy selected for services with particular characteristics.
Generally, selecting a suitable combination of links and nodes in a network architecture supporting a p-mp connection is complexer than selecting a p-p route [WAX88, TAN94, DOAR93]. Obviously, the routing of mp-mp connections is even more complicated by the additional requirements of returns paths in the connection pattern [VERM93]. Implementation of routing procedures responsible for selecting suitable routes requires therefore smart routing algorithms.
According to the conclusions of the previous chapter, multiple destination routing algorithms must suit different routing strategies. State-dependent routing requires a near real-time control mechanism based on current network status information. Time-dependent routing requires adaptability to recurrent network status patterns and thus requires long term information. On the other hand, mutative routing puts totally different requirements on routing algorithms: connections grow and shrink in a gradual manner.

Another basic requirement is formed by the non-symmetrical character of p-mp connections. Network link load imbalances must be avoided by a mechanism supported by the implemented routing algorithm.
Important is to recognize how routing algorithms to be implemented maximize interoperability, as this is one of the main objectives of network design. Thus, interworking between routing systems for p-p connections, p-mp connections and mp-mp connections should be endeavoured, whether they carry unidirectional, bidirectional or bidirectional asymmetric traffic streams.

Research in multiple destination routing has been limited. Results contributed employ mostly very specific p-mp routing problems or put impractical constraints on the implementation in telecommunication networks. However, available algorithms are collected and will be described in the next sections providing useful concepts for the design of routing systems. Some algorithms are contributed from the mathematical branch of graph theory, some others from (computer) network designers. The main objective is to classify the algorithms according to previous described requirements on routing. The classification of MDR algorithms will then provide links between algorithms and network implementation which are needed to make a judgement.

4.1.2 Multiple Destination Routing Problem

The complexity of multiple destination routing problems requires a rather mathematical approach. Thus, a mathematical model of the network is needed. A telecommunication network architecture consists of bidirectional links and exchanges between those links, whether it concerns a physical or a functional architecture. Using mathematical graph theory, the network can be modeled as a directed graph $G = (N, L)$, where N is a set of nodes representing network exchanges and L is a set of directed links l representing the network links. A bidirectional network link between nodes $u, v \in N$ is thus represented by two directed links in the graph G denoted by (u, v) and (v, u).
A number of available MDR algorithms operate only on undirected graphs. Modeling the network as an undirected graph presumes equal desirability for both link directions to include them in the selected route. In this case, a particular undirected graph link is again denoted by $l \in L$ but now it represents a complete bidirectional network link.

The routing algorithms to be implemented in the network must fit within the routing strategy in use. Therefore, the routing strategy must provide criterias to the routing algorithms determining the desirability of using a particular network link. These criterias are modeled by a link cost function C giving each network link a positive link cost $c(l)$ indicating desirability of using link l. The parameters contained in this cost function can be the link utilization ratio, the ratio of imbalanced link load or the delay associated with a link for example. Generally, these parameters must represent the short term network status in case of state-dependent routing or represent long term utilization patterns in case of time-dependent routing.

A multiple destination routing problem is the problem of finding a routing R between the *source* node $s \in N$ and the set of *destination* nodes $D \subset N$, with $s \notin D$ such that the accumulated link costs of links included in the route is minimized. The accumulated costs of a routing R will indicated by *network cost* NC(R).

The solution of an MDR problem consists of a *routing tree* which is a rooted subtree of the graph (N, L) whose root is s, that contains all of the nodes of D, and an arbitrary subset of $(N - D)$, and whose leaf set consists only of a subset of nodes of D.

The intuitive interpretation of a routing tree or routing is that node s sends a copy of the information cells (to be sent to D) to each son of s in the tree. These sons in turn transmit the information cells to their sons until all nodes in the tree (and thus all nodes in D) have received the information cells.

The above described problem of finding an optimal routing tree by minimizing link costs is a well-known problem in the graph theory which is called the *Minimum Steiner Tree* problem [KOU81, BHA83, WIN87, JIA92]. A *Steiner tree* is a spanning tree connecting a subset of node set N. A Minimum Steiner Tree connects (spans) a subset of node set N at minimal accumulated link cost . The problem of finding that tree belongs to the class Non-deterministic Polynomial-complete (NP-complete). This means that the required computation time for finding a minimum Steiner tree increases exponentially with increasing network size. Therefore, heuristic algorithms are developed using polynomial bounded computation times for finding near-optimal solutions.

4.1.3 Criterias for classification of MDR algorithms

A number of algorithms are available which can be used to generate solutions for particular MDR problems. However, different routing strategies lay distinct constraints on routing algorithms. Generally these constraints relate to:

- availability of network status information
- centralized or distributed implementation of routing control.

Considering the point of availability of network status information, a basic distinction exists between *global information* algorithms and *non-global information* algorithms. Global information is defined by status information of the complete network in consideration. On the other hand, non-global information is status information of certain part of the network. Within the class of non-global information algorithms, the amount of available information can vary, resulting in:

- nearby information
- local information
- strict local information.

This will be explained together with the second criterion which demarcates MDR algorithms. A routing algorithm either runs on one network node having global information or runs in a distributed fashion supported by the whole network i.e. *centralized control* or *distributed control.*

In a centralized routing control, a set of special nodes in the network make routing decisions. In principle, these special nodes periodically receive current status information such as cost, traffic and congestion data from all the other nodes and uses this information to determine routing trees. Thus, protocols are required for the nodes in the network to send updating information to the control nodes. In addition, protocols are required for the control nodes to send its routing decisions to the other nodes.

In a distributed routing control, there are no special nodes which make the routing decisions. Each node in the network can have the responsibility to determine a complete routing tree or a part of the routing tree.

Distributed routing control can be further differentiated into intermediate-node routing and source-node routing [JAF85]. Source-node routing is when each node in the network determines the exact route of which it is the root. Therefore, source-node routing requires global network information. Intermediate-node routing is when each node determines only the next link(s) in the route. Thus, intermediate-node routing requires typically less network information than source-node routing.

Source-node routing requires each node of the network to maintain an up to date topology database containing all link costs of the network. In contrast, intermediate routing needs only non-global information. Intermediate-node routing algorithms require one of the above enumerated types of information. The first, *local information*, is defined to be the accumulated costs of the shortest paths from the node in consideration to all other network nodes and the knowledge of neighbouring nodes on these shortest paths. The shortest path is defined here as the route between two network nodes with minimal accumulated link costs. In addition to this, *nearby information* is defined as the local information of the neighbouring nodes.
Strict local information is defined as the knowledge of the capacity of links and the cost of links to which a given node is connected. Figure 15 shows the relationships between distinct information and control criterias which characterize MDR algorithms.

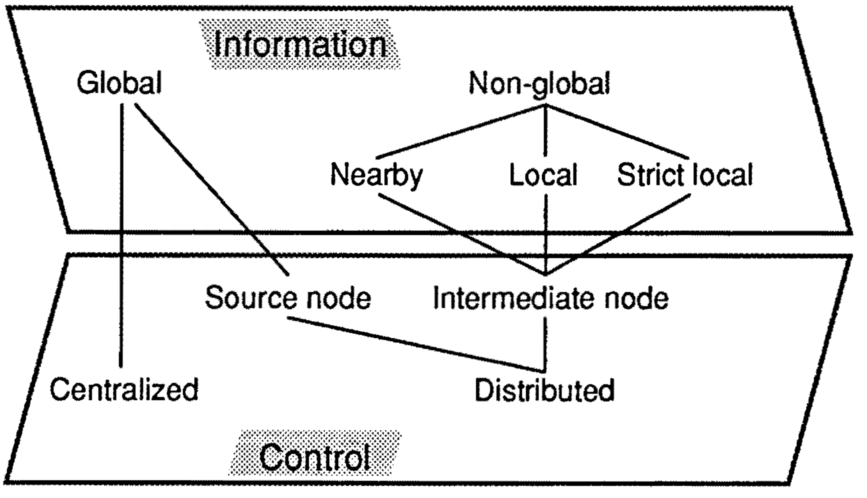


***Figure 15** Relationships between the several criterias*

Mutative routing

Multiple destination routing algorithms can also be categorized according to their value for mutative routing strategies. As described in the previous chapter, requirements on mutative routing originate from situations where new subscribers request connection to a communication session in operation, or where current subscribers request termination. Given this, it is necessary at the very least to establish routes for the new subscribers. In ATM networks, rearrangement of already established connections is not allowed to maintain cell sequence integrity at the destinations.

4.2 Available global information MDR algorithms

Four heuristic algorithms will be discussed in this section that use global information and attempt to construct minimal Steiner trees. The Minimum Spanning Tree heuristic and the Rayward-Smith heuristic are both non-mutative approximation algorithms. The Weighted Greedy algorithm is appropriate for mutative routing. The algorithms described here can be used either for centralized control or source-node routing. The first three global information algorithms model a telecommunication network as an *un*directed graph, which implies that the cost of using a particular link is equal for both the communication directions. The last algorithm described in this section uses a directed graph and originates from linear programming theory.

4.2.1 The Minimum Spanning Tree (MST) heuristic

The MST heuristic is a well-known algorithm for the Steiner tree problem, the original form of this algorithm is described in [KOU81]. The latest improvement for faster execution can be found in [MEHL88].

The MST heuristic operates on undirected connected graphs. Because there exists no distinction between source node and destination nodes in undirected routings, source node s and destination nodes D are combined in $S = \{s\} \cup D$. The 'bridging' nodes in the minimum spanning tree of S but not contained in S are called *Steiner nodes*. The MST heuristic tries to find a minimum spanning tree in graph $G = (N, L)$ for node set S in five subsequent steps (see also the example in figure 16).
The first step constructs a new subgraph $G_1=(S, L_1)$ of G containing only the nodes to be connected (S) and a new link set L_1 representing the shortest paths in G between those nodes S. Because this first step is most time consuming of the whole algorithm, Mehlhorn used a special smart method for this step by defining specific partitions of the graph for each node in D. These partitions fasten the search process for the shortest paths.
Step two constructs a minimum spanning tree G_2 of the new (simpler) graph. This spanning tree is then used to select the corresponding links and (additional) nodes in the original graph G resulting in graph G_3. Step 4 constructs again a minimum spanning tree G_4 by deleting links with the highest costs from G_3. The last step deletes superfluous edges, so that no leaves in the resulting solution R are Steiner nodes.

The exact mathematical formulation is as follows. Some definitions are needed first.
A path in G is a sequence of links $l_1, l_2,\ldots, l_k$ of N such that, for all i, $1 \le i < k$, $(l_i, l_{i+1}) \in L$ is an edge of the graph. The cost of a path is the sum of the costs of the edges. The cost of the shortest path between two nodes u, v in G is denoted by $mc(u, v)$. For every node $d \in S$ let $M(d)$ be the set of nodes in N which are not further from d than from any other node in S. More precisely, consider a partition $\{M(d); d \in S\}$ of N, i.e. $\cup_{d \in S} M(d) = N$ and $M(d) \cap M(e) = \varnothing$ for $d, e \in S$, $d \neq e$, with

$$v \in M(d) \Rightarrow mc(v, d) \le mc(v, e) \text{ for all } e \in S.$$

The MST heuristic as proposed in [KOU81] and modified in [MEHL88] is described as follows:
1. Construct a new subgraph $G_1=(S, L_1)$ of G defined by:

$L_1 = \{(d, e); d, e \in S$ and there is a link $(u, v) \in L$ with $u \in M(d), v \in M(e)\}$
and
$mc_1(d, e) = \min\{mc(d, u) + c(u, v) + mc(v, e); (u, v) \in L, u \in M(d), v \in M(e)\}$.

2. Find a minimum spanning tree G_2 of G_1.
3. Construct a subgraph G_3 of G by replacing each link in G_2 by its corresponding shortest path in G. (If there are several, pick an arbitrary one.)
4. Find a minimum spanning tree G_4 of G_3.
5. Construct a solution R from G_4 by deleting edges in G_4, if necessary, so that no leaves in R are Steiner nodes.

The theoretical worst case performance is usually expressed as the ratio of the network cost of the generated solution to the network cost in the optimal case. For MST algorithm, it is known that $NC(R_{MST})/NC(R_{opt}) \leq 2[1-(1/t)]$, where an optimal tree has t leaves [HWA92]. However, simulation results show that the average performance is very close to optimal: ratio to optimal $\leq$ 1.05 [TAN93].
The complexity analysis of the algorithm is given in [MEHL88]. It appears that step 1 is the most time consuming step and takes $O(|S| \log|S| + |L|)$ time.

Example
$G = (N, L)$ is given in figure 16a. Each link is labelled with an integer which denotes the cost. Let $S = \{v_1, v_2, v_3, v_4\}$ be the set of nodes to be connected. Figure 16b shows the graph G_1. Figure 16c shows the minimal spanning tree G_2 of G_1. Figure 16d shows the corresponding subgraph, G_3 of G. The minimal spanning tree G_4 of G_3 is shown in figure 16e. Figure 16f shows the final output for this example, R. This example demonstrates the fact that G_2, G_3 and G_4 might not be unique.

4.2.2 The Rayward-Smith heuristic

The Rayward-Smith (RS) algorithm proposed by Rayward-Smith in [RAY86] attempts to find Steiner nodes which could be useful for the minimum spanning tree. An additional requirement is that the node must have a degree of three or more (the degree deg(s) of a node s is the number of edges which have s as an endpoint). Rayward-Smith defines a function $f: N \rightarrow \Re^+$ such that those nodes for which the function has a minimum value are the ones that should be included in the final solution. The algorithm is described as follows:

1. Construct a collection of single node trees $\mathfrak{I}$ consisting of the nodes in $S = \{s\} \cup D$
2. Repeat the following steps until $|\mathfrak{I}| = 1$.
   a. choose $n \in N$ such that $f(n)$ is minimum where

$$f(n) = \min_{U \subseteq \mathfrak{I}, |U|>1} \{\frac{1}{|U|-1} \sum_{T \in U} \mathrm{dist}(n,T)\} \tag{4.1}$$

   f is the average cost of joining r to $r + 1$ trees through a node n where $r + 1 = S$ and dist (n, T) is the distance (cost) between node n and T.
   b. Let T_1 and T_2 be the closest trees to n.
   c. Join T_1 and T_2 by a shortest path through n.

Example
The application of RS to an eight node graph is illustrated in figure 17. In this example, the repeat statement is executed four times to find a minimum spanning tree for the nodes $\{v_1,v_3,v_4,v_5,v_7\}$.

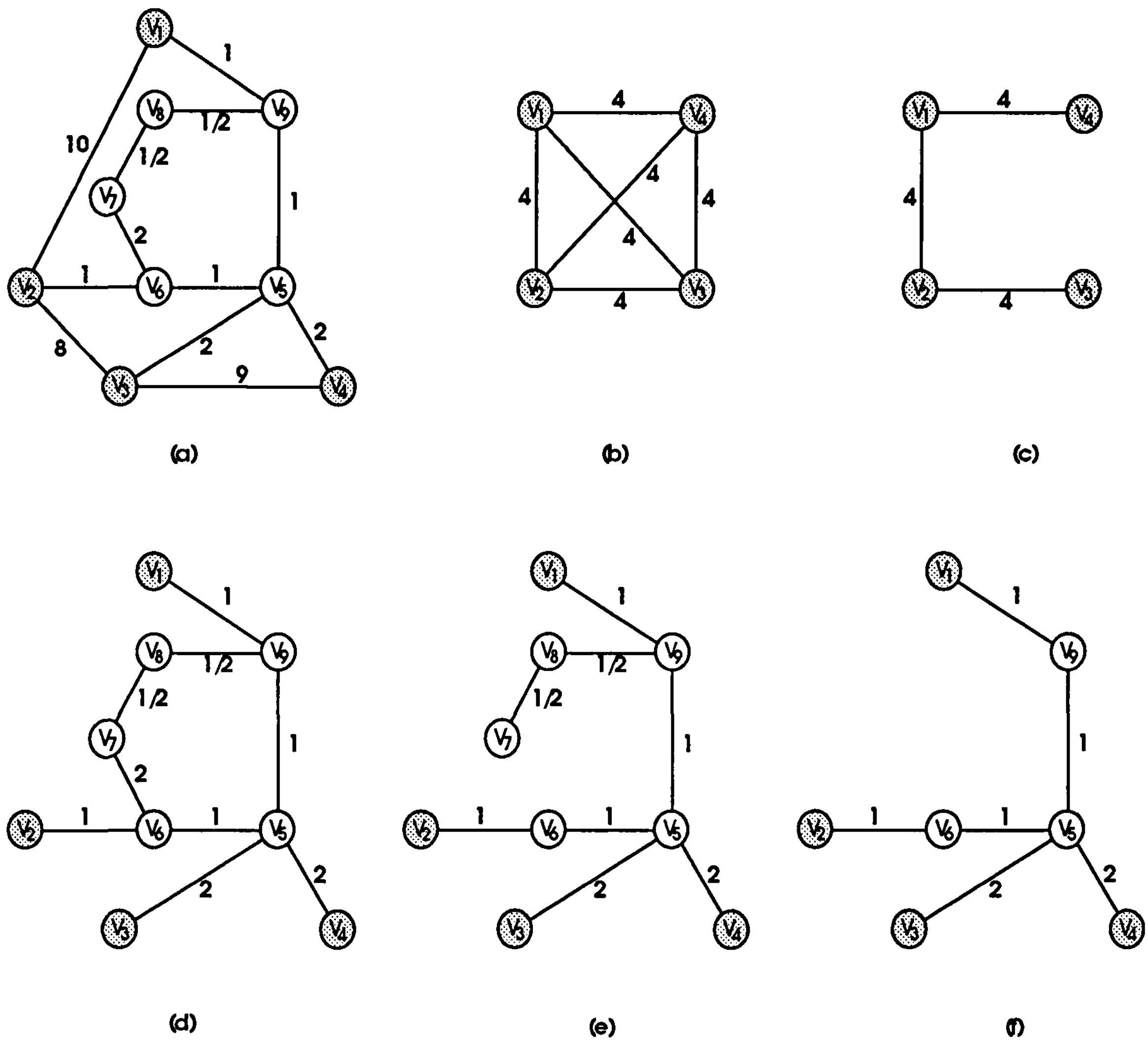


Figure 16 Example of an application of the MST algorithm to a 9-node graph

The worst case bound on the performance ratio of the RS algorithm is claimed to be within two times optimal [HWA92]. From [WAX88], it can be concluded that the average performance ratio to optimal is typical below 1.01. Considering the time complexity, the algorithm runs in $O(|S|^3)$ time.

4.2.3 Weighted Greedy algorithm

The Weighted Greedy algorithm (WGA) is a quasi-static algorithm that considers a sequence of requests $R=[r_0,\ldots,r_n]$, which apply to either an *add* node request or a *remove* node request [WAX93, DOAR93]. There is always a *connection owner*, which is the source node in the routing tree. The problem is to find the sequence of routes $P=[p_1,\ldots,p_2]$ which yields a minimum cost sequence. In the case of mutative routing, once a particular set of links has been used in a route, no rearrangement of these edges is allowed as the algorithm proceeds.
The Weighted Greedy algorithm is a more general variation of the Greedy algorithm. In the Greedy algorithm, to add a node u to an existing connection, the closest node v already in the connection has to be found. Node u is then connected to v by the shortest path to v. To delete a node from the connection, it is first marked as 'deleted'. If that node is not an internal node in

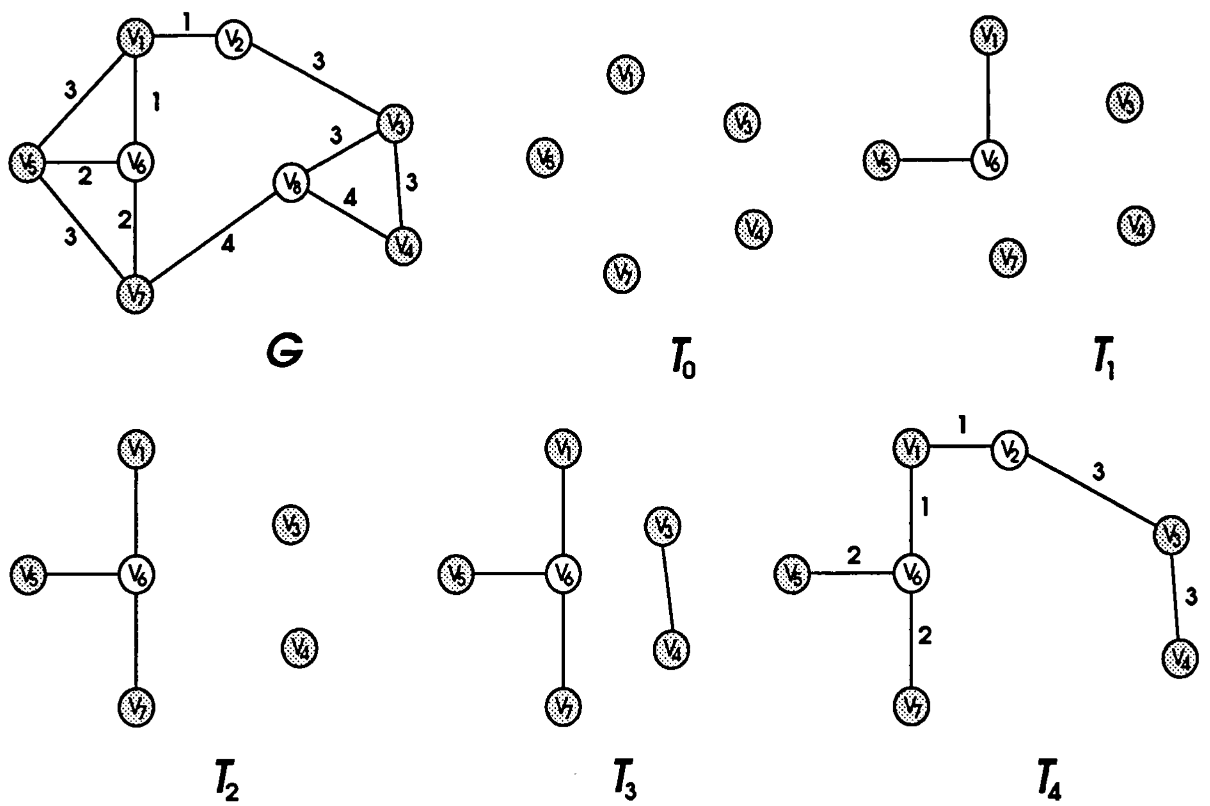


Figure 17 Example of application of the RS algorithm to an 8-node graph

the connection, the branch of which it is a part is pruned from the connection.
The WGA also joins a new node through a path to the existing connection tree $T=\{N(T), L(T)\}$. But in this case, the path is selected to minimize a function of the distance to T and the distance within T, to the owner v_0. To add a node u to a connection, a node $v \in N(T)$ is selected to minimize the function

$$W(u,v) = (1-\omega)d(u,v) + \omega d(v,v_0) \tag{4.2}$$

where $0 \le \omega \le 0.5$ and $d(x,y)$ is the distance (or cost) between a pair of nodes x and y. The parameter ω determines the weight given to the distance from u to v relative to the weight given to the distance from v to the connection owner v_0. If $\omega = 0$ only the distance from u to v is considered and is equivalent to the greedy algorithm. Thus a shortest path to the connection is selected. When $\omega = 0.5$, the distance from u to v and the distance from v to v_0 are given equal weight. In this case a shortest path to the owner is selected. When $0<\omega<0.5$, the algorithm connects the new node v to the connection T through a path that points towards source node v_0 but not necessarily the nearest path to that source node. In this manner, the connection pattern forms a tree-like pattern whose origin is the source node v_0. This algorithm assumes that as users are added and removed, the connection pattern will not have a loop, but remains in a tree-like pattern pointed toward the origin.
The time complexity of the WGA depends on the implemented shortest path routing algorithm. The performance of the WGA was investigated in [DOAR93]. Simulations showed that the cost of solutions generated by the WGA are typical within a factor of 1.5 compared to the cost of routing trees generated with the MST algorithm.

4.2.4 Dual Ascent Algorithm

The Dual Ascent Algorithm (DAA) is a linear programming approach to the Steiner problem on *directed* graphs [WONG84]. The problem is to find a minimum cost directed subtree that contains a directed path between the source node and each destination node. Thus graph $G = (N, L)$ represents in this case a node set N connected by link set L containing directed links (also called arcs). A directed link from node i to node j is indicated with (i, j). The Steiner tree problem on a directed graph is represented using an integer programming formulation:

$$(P_1) \quad \text{minimize} \quad \sum_{(i,j)\in L} c_{ij} y_{ij} \tag{4.3}$$

$$\sum_{h\in N} x_{ih}^{k} - \sum_{j\in N} x_{ji}^{k} = \begin{cases} 1, & i=1, \\ -1, & i=k, \\ 0, & i\neq 1,k, \end{cases} \quad k\in D \tag{4.4}$$

$$x_{ij}^{k} \leq y_{ij}, \tag{4.5}$$

$$x_{ij}^{k} \geq 0,\ (i,j)\in L,\ k\in D,$$

$$y_{ij} = \in \{0,1\} \tag{4.7}$$

where y_{ij} is a binary variable indicating whether or not link (i,j) is included in the solution. The variable x_{ij}^{k} is the amount of commodity k (the amount of 'multicast flow' between the source node 1 and k) on link (i,j). Constraint (4.4), (4.5), (4.7) indicates that flow on a link is allowed only if the link is included in the solution. Constraint (4.4), (4.5), (4.7) dictates that one unit of commodity k be routed between nodes 1 (source) and k. In constraint (4.4), (4.5), (4.7) y_{ij} is viewed as the flow capacity for commodity k on link (i,j). Constraints (4.4), (4.5), (4.7)-
(4.4), (4.5), (4.7) indicate that a feasible solution must have a directed path of links (i.e. $y_{ij} = 1$) between node 1 and every destination node. Thus the connectivity of the Steiner tree problem is represented via an embedded multi-commodity network flow problem.
The solution for problem P_1 is first approximated by considering the linear programming relaxation LP_1 of P_1 and its dual DLP_1, at the same time the lower bound for the solution to P_1 is obtained. The dual of LP_1 can be written as:

$$(DLP_1) \quad \text{maximize} \quad \sum_{k\in D} (v_k^{k} - v_1^{k}), \tag{4.8}$$

$$v_j^{k} - v_i^{k} - w_{ij}^{k} \leq 0,\ k\in D,\ i,j\in N \tag{4.9}$$

$$\sum_{k\in D} w_{ij}^{k} \leq c_{ij},\quad (i,j)\in L, \tag{4.10}$$

$$w_{ij}^{k} \geq 0. \tag{4.11}$$

The dual variable v_i^k is associated with the conservation of flow equation for commodity k at node i. As is usually the case for a complete set of flow balance equations, one equation is redundant, so $v_1^k = 0$ for all $k \in D$.
Before the algorithm is described, some definitions are given first. An auxiliary directed graph $G' = (N, A)$ is defined, which has the same node set as G but has a special set of links, a subset of the original set of links L, which will be specified later. Node i is *connected* to node j if there

is a directed path from i to j in the auxiliary graph G'. Nodes i en j are *strongly connected* if i is connected to j and vice-versa. A set of nodes Q is strongly connected if any two nodes in Q are strongly connected. A maximal strongly connected set is a *strongly connected component* of G'. If node i is connected to node j but not vice-versa, then node i *dangles* from node j. A strongly connected component T is also a *root component* if T contains a member of D but no member of $D \cup s$ dangles from a member of T. The complement of any set T is denoted by T^-. Let $C(k)$ be the set of nodes that are connected to node k. The *cut set* of k is a set of arcs whose members (i, j) satisfy (1) $(i, j) \in L$; (2) $(i,j) \notin A$; (3) $j \in C(k)$ and $i \in C^-(k)$. In other words, arc (i, j) is a potential member of A which would connect node i to node k.
The ascent procedure for computing a near optimal solution to DLP_1 is as follows.

Step 0. Initialize:

$$v_i^k := 0, \quad k \in D,\ i \in N,$$
$$w_{ij}^k := 0, \quad (i, j) \in L,$$
$$S(i, j) := c_{ij} - \sum_{k \in D} w_{ij}^k = c_{ij}.$$

Form the auxiliary graph $G' = (N, A)$ with $A = \emptyset$. Since $A = \emptyset$ initially all nodes are strongly connected components and all members of D are root components.
Step 1. Select a root component R. If there are no root components then **STOP**.
Step 2. Select a node $k \in D \cap R$. Let $S(i^*, j^*) = \min\{S(i, j) \mid (i, j) \in \text{cut set of } k\}$.
For each node $h \in C(k)$,

$$v_h^k := v_h^k + S(i^*, j^*).$$

For each $(i, j) \in$ cut set of k

$$w_{ij}^k := w_{ij}^k + S(i^*, j^*).$$

and

$$S(i, j) := S(i, j) - S(i^*, j^*).$$

Step 3. Update the auxiliary graph by setting

$$A := A \cup \{(i^*, j^*)\}.$$

Go back to Step 1.

It has been proven in [WONG84] that this procedure generates the optimal solution for the directed spanning tree problem when all nodes have to be linked (this is the case when $s \cup D = N$).

The overall procedure for the directed Steiner problem has the following description:

(1) Use the dual ascent procedure as described above to obtain a lower bound for the optimal value;

(2) Consider the auxiliary graph G' when the ascent algorithm terminates and let Q be the set of all nodes connected to the root node. The root node and all members of D will belong to Q (the algorithm will not terminate unless all members of D are connected to the root).

(3) Form a minimal directed spanning tree problem with node set Q. The arc set will be any arc $(i, j) \in L$ where i and $j \in Q$. The arc costs c_{ij} are taken from the original graph. Solve the directed spanning tree problem with the ascent algorithm and a procedure for recovering the optimal tree from the auxiliary graph G'.

(4) Let T be the minimal spanning tree found in Step 3. If node $n \notin D$ is a leaf of T delete node n and the arc incident to it. Repeat this process until no further changes are required.

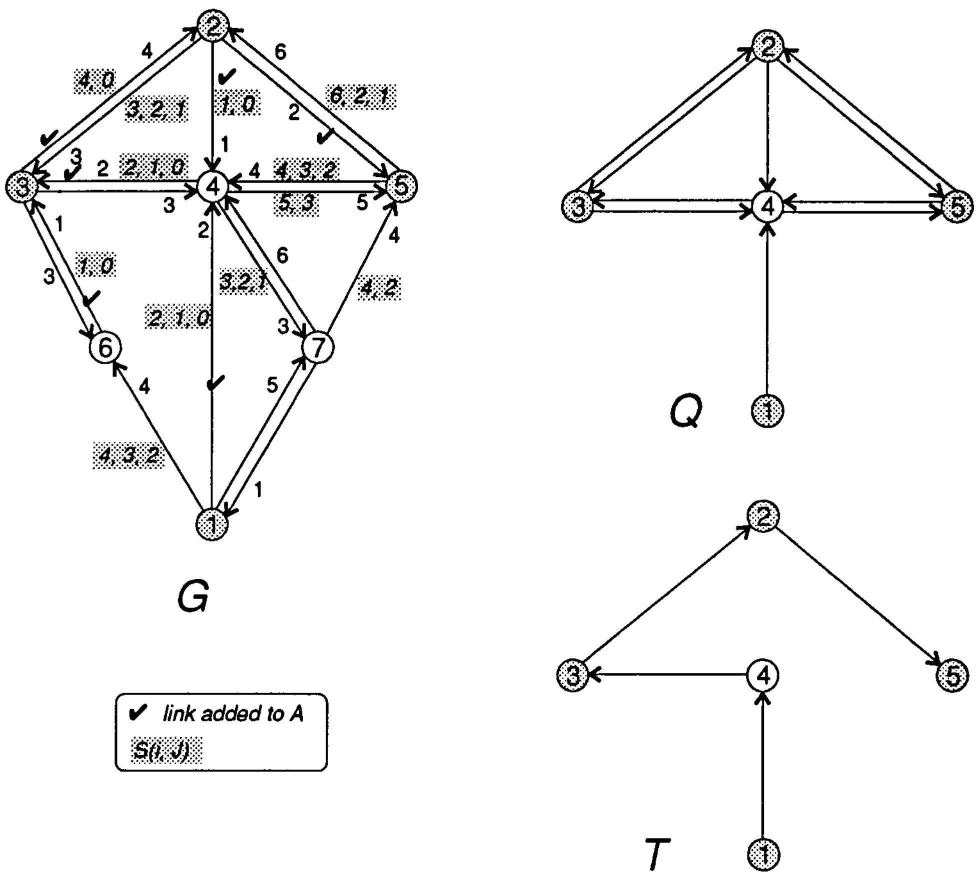


***Figure 18** Example of application of the Dual Ascent algorithm*

Example

The application of the Dual Ascent Algorithm to an eight node directed graph is illustrated in figure 17, 18. The problem is to find a minimal multicast tree from source node 1 to destinations {2, 3, 4}. Table 4.1 details step one of the overall procedure. The dual ascent algorithm terminates after six iterations, the dual objective function value is

$$\sum_{k \in D} v_k^k = 5 + 3 + 2 = 10$$

This is the lower bound for the optimal solution.

Table 4.1 Dual Ascent Algorithm for the example of figure 18

Iteration	R	k	Cut set of k	$[i^*, j^*, S(i^*, j^*)]$	Variables modified
1	2	2	(3, 2)(5, 2)	[3, 2, 4]	$v_2^2 = 4 = w_{32}^2 = w_{52}^2$
2	3	3	(2, 3)(4, 3)(6, 3)	[6, 3, 1]	$v_3^3 = 1 = w_{23}^3 = w_{43}^3$
3	5	5	(2, 5)(4, 5)(7,5)	[2, 5, 2]	$v_5^5 = 2 = w_{45}^5 = w_{75}^5$
4	3	3	(2, 3)(1, 6)(4, 3)	[4, 3, 1]	$v_3^3 = 2 = w_{43}^3 = w_{23}^3$, $w_{16}^3 = 1$
5	3	3	(1, 6)(1, 4)(7, 4) (5, 4)(2, 4)(2, 3)	[2, 4, 1]	$v_3^3 = 3 = w_{23}^3$, $w_{14}^3 = w_{24}^3 = w_{54}^3 = 1$, $w_{74}^3 = 1$, $w_{16}^3 = 2$
6	{2,3}	2	(1, 6)(1, 4)(7,4) (5, 4)(5, 2)	[1, 4, 1]	$v_2^2 = 5 = w_{52}^2$ $w_{16}^2 = 1$ $w_{14}^2 = w_{54}^2 = w_{74}^2 = 1$

Computational results

In [WONG84] results of computational experiences with the dual ascent algorithm are given. The algorithm resulted 22 times in an optimal solution for graphs containing up to 60 nodes and 240 directed arcs and 2 times in an non-optimal solution with an upper bound/lower bound ratio less than 1.01. These results indicate that the algorithm is very effective in solving minimum spanning tree problems in directed graphs.

4.2.5 Other global information algorithms

Other centralized control, global information, static algorithms have been proposed which take one special network constraint into account: a restricted number of copy operations in intermediate nodes in the multicast tree. They are the (modified) link-add-type and (modified) loop-construct-type algorithms [TAN93]. Simulation results which give an indication of their performance under diverse network load circumstances are not available.

4.3 Available non-global information algorithms

This section describes three non-global, intermediate-node approximation algorithms which have been proposed for multiple destination routing in computer networks [BHA83]. The philosophy behind these algorithms was to design 'practical' algorithms for minimum cost routing by using local shortest path information. As defined earlier, local information consists of the accumulated costs of reaching all other network nodes from the node in consideration via shortest paths. The non-global information algorithms described here can be applied to both directed and undirected graphs by making distinction between shortest paths for forward and backward information transport. For practical reasons, the illustrating examples used here are generalizations to undirected graphs.

Intermediate-node routing algorithms (also called Forward Call Control algorithms) generally forward the responsibility for further setting up the connection to neighbouring nodes which are selected by the current decision making node to make part of the route. The process is initiated at the source node and ends when all destination nodes are reached. Each node makes routing decisions based on its own network information.

Intermediate-node routing algorithms use network status information applicable for the routing of all types of connections (p-p, p-mp, mp-mp).

4.3.1 Minimum Destination cost algorithm

The first algorithm is called the MINimum Destination Cost (MINDC) algorithm because it uses only local shortest path information and minimizes for that very reason *destination cost* and not accumulated costs of links contained in the resulting routing tree. Destination cost (DC) of a routing R is defined by

$$\mathrm{DC}\,(R) = \frac{1}{|D|}\Big[\sum_{d \in D} \sum_{l \in p(d)} c(l)\Big] \tag{4.18}$$

where $p(d)$ is the path in R from s to d and $|D|$ is the number of destination nodes. If $c(l)$ were the delay of l, then DC(R) would represent the average delay experienced by the destinations. Thus, the MINDC algorithm minimizes cost by choosing the shortest path for each pair of source to destination.

MINDC

```
function ForwardConnect (D_current)
define:
cn = current node
neighbour(x) = neighbouring node on shortest path to destination x
begin
  if cn ∈ D_current then D_current := D_current - cn
  if D_current ≠ ∅ then begin
    while D_current ≠ ∅ do
    begin
      GroupToForward := ∅
      choose arbitrary d_1 ∈ D_current
      D_current := D_current - d_1
      connect neighbour(d_1) to routing tree
      add d_1 to GroupToForward
      for all remaining d ∈ D_current do
        if neighbour(d) = neighbour(d_1) then add d to GroupToForward
      ForwardConnect(GroupToForward)
      D_current := D_current - GroupToForward
    end
  end
end {ForwardConnect}

Main
begin
  start at source node
  ForwardConnect(D)
end.
```

The theoretical worst case performance for MINDC is n times worse than optimal, where n is number of nodes in the network [BHA83].

Example

Figure 19 shows the resulting routing tree generated by the MINDC algorithm for two distinct source/destinations sets. The connections are constructed with the shortest paths between source and destinations resulting in total a cost of 13 for both connections.

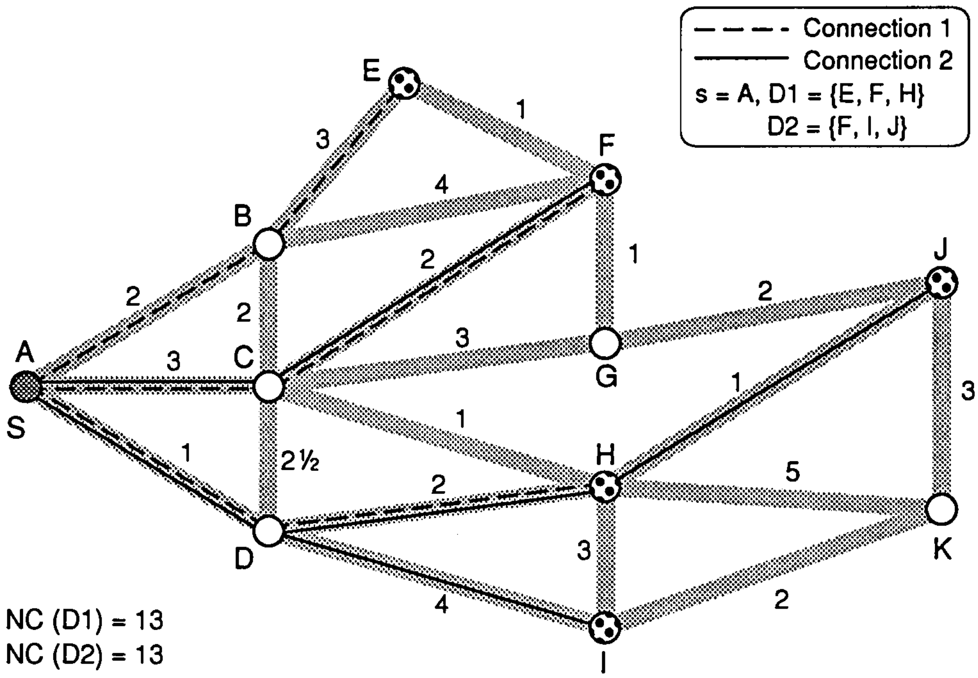


***Figure 19** Resulting routing trees with the Minimum Destination Cost heuristic*

4.3.2 Nearby Heuristic 1

The second algorithm Nearby Heuristic 1 (NH1) is an intermediate node algorithm that uses nearby information, i.e. all information available to its surrounding neighbours, as well as its own local information. The NH1 operate as follows.
Assume that s were to bifurcate to neighbours n_1, n_2, n_3. Assume further that s routes to a set of destinations D_i through n_i with $D_i \cap D_j = \emptyset$ for $i \neq j$ and $\cup_i D_i = D$. Then the *local cost* of the routing is the sum of the costs of the three links, connecting s to n_1, n_2 and n_3. The *estimated global cost* is the sum of the costs from n_1 to each destination node in D_1, plus the sum from n_2 to each node in D_2, plus the sum from n_3 to each node in D_3. Roughly speaking, the estimated cost assumes that each destination is routed to separately after n_1, n_2 and n_3. The *total estimated cost* for the routing is the sum of the local cost and the estimated global cost. Algorithm NH1 routes from s to its neighbours by choosing the routing that has the smallest total estimated cost. Once this is done, intermediate nodes choose their next nodes in the same way.

Nearby Heuristic 1

```
function ForwardConnect (D_current)
define:
cn = current node
mc(x, y) = cost of shortest path between nodes x and y
n_1,..n_j neighbours of cn except originating node
D_1,..D_j sets of destination nodes to be assigned to n_1,..n_j
begin {ForwardConnect}
  if cn ∈ D_current then D_current := D_current - cn
  if D_current ≠ ∅ then begin
    consider each possible combination (n_1, D_1)..(n_j, D_j)
    compute TotalEstimatedCost =
```

$$\sum_{1 \le i \le j} e_i c(cn, n_i) + \sum_{d_i \in D} mc(n(d_i), d_i)$$

```
        where e_i is a boolean indicating whether D_i is empty
    select the combination with lowest TotalEstimatedCost
    for 1 ≤ i ≤ j do
      if e_i then begin
        connect n_i
        ForwardConnect (D_i)
      end
  end
end {ForwardConnect}

Main
begin
  start at source node
  ForwardConnect(D)
end.
```

The computation time requirement of this heuristic grows exponentially with number of neighbours.

Example

Figure 20 illustrates the result of Nearby Heuristic 1 applied to the same source/destination sets as the previous example. Considering connection 1, the procedure starts in source node A in order to connect destinations E, F and H. The following combinations of neighbours with destinations are possible (neighbour, {destinations}):

(B, {E, F, H}); cost = 12
(C, {E, F, H}); cost = 9
(D, {E, F, H}); cost = 13
(B, {E} + C, {F, H}); cost = 12
(B, {E, F} + D, {H}); cost = 12
(C, {E, F} + D, {H}); cost = 12
(B, {E} + C, {F} + D, {h}); cost 13

The second combination is the cheapest, thus node C is linked first and it is simultaneously charged for connecting destinations E, F, H. The procedure running at node C clearly selects now destination nodes F and H as the next nodes to be linked. Node F completes the routing tree by connecting destination E. The total cost is of connection 1 is 3+2+1+1=7.

This example shows that using additional nearby information results in stronger convergence in the top of the routing tree compared with the results of the MINDC algorithm.

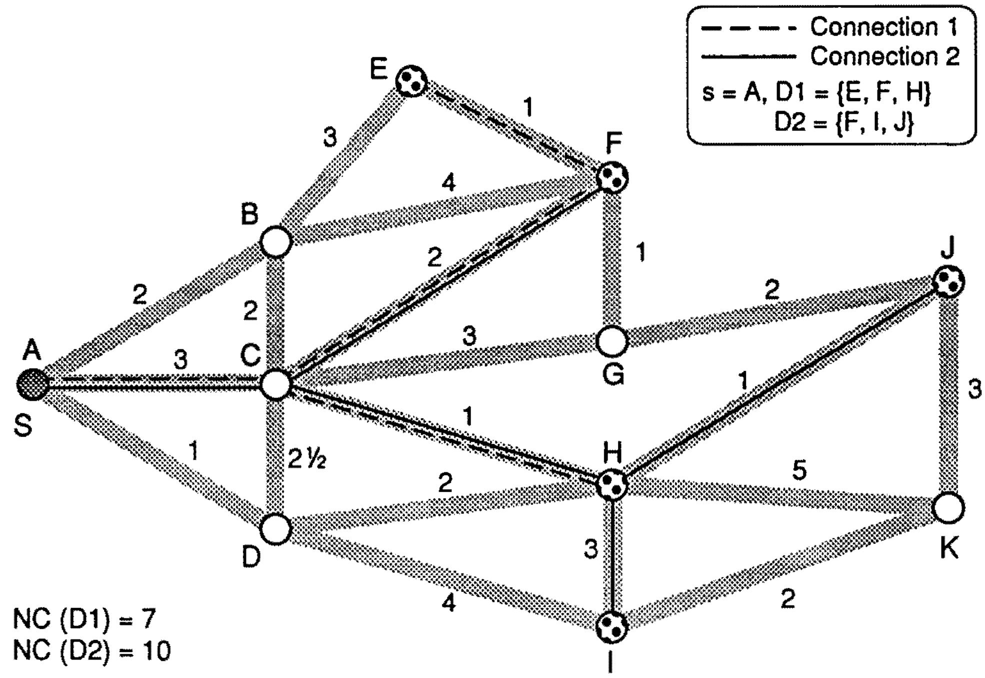


***Figure 20** Resulting routing trees of Nearby Heuristic 1*

4.3.3 Nearby Heuristic 2

Nearby Heuristic 2 (NH2) investigates the possibility of using the neighbours one by one, rather than trying all possibilities of neighbours. The destination d closest to s is chosen first, and the neighbour n_1 which lies on the shortest path from s is investigated. For each $d' \in D$, if the least cost path from n_1 to d' is smaller than that from s to d', d' is assigned to n_1. After every d' is examined, if any are left unassigned, the one that is closest is chosen next. Let n_2 be the neighbour on the shortest path from s to it. The algorithm now attempts to assign the remaining destinations to n_2 as above. The procedure continues until all destinations are assigned to neighbours of s. Then, the analogous procedure is used at each of these neighbours.

Nearby Heuristic 2

```
function ForwardConnect (D_current)
define:
cn = current node
mc(x, y) = cost of shortest path between nodes x and y
begin {ForwardConnect}
 if cn ∈ D_current then D_current:= D_current - cn
 while D_current ≠ Ø do
 begin
  select d closest to cn
  connect neighbour n' on shortest path to d; D_current:= D_current - d
  GroupToForward:= d
  for each remaining d in D_current do
  begin
   if mc(n', d) ≤ mc(cn, d) then add d to GroupToForward
   D_current:= D_current - d
  end
  ForwardConnect(GroupToForward)
 end {while}
```

```
end {ForwardConnect}

Main
begin
  start at source node
  ForwardConnect(D)
end.
```

Example

Figure 21 illustrates again routing trees for the same source/destinations combinations from the previous examples generated by Nearby Heuristic 2. Considering destination set 1, node H is closest to A, thus neighbouring node D on the shortest path to H is added first to the connection. The shortest path from node D to F (4½) is shorter than the shortest path from source node A to F (5), so nodes F and H are assigned to D; destination E is assigned to node B. The procedure at node D selects then node H as the next node to connect. The shortest path from node H to F is shorter than the shortest path from D to F, so destination F is assigned to node H.
The total cost of connection 1 is 12, the total cost connection 2 is 10.

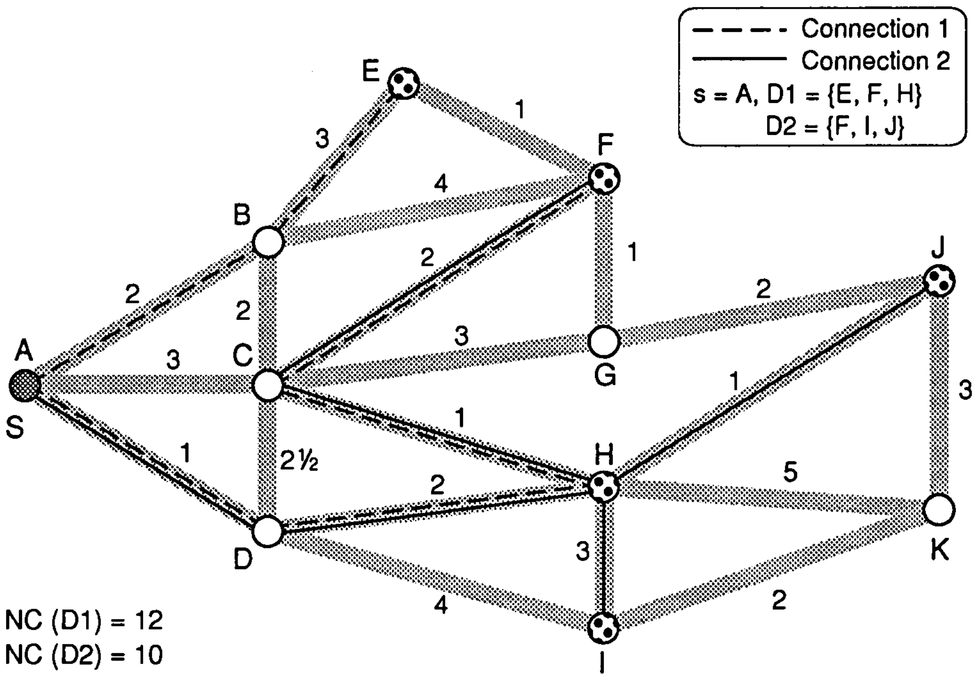


Figure 21 Resulting routing trees of Nearby Heuristic 2

4.3.4 Flooding algorithms

Flooding algorithms operate in a distributed fashion and their main purpose relates strongly to maintaining network status information at specified points in the network (e.g. at all network nodes). However, flooding techniques can also be used for making connection specific routing decisions as described by X. Jiang [JIA93].
Flooding algorithms require only strict local information. In fact, local knowledge of neighbouring nodes and link status information of links connecting those neighbours is sufficient to make routing decisions or to compose shortest paths between given node pairs. The

basic concept of flooding techniques is the iterative broadcast of connection request messages throughout the whole network in consideration. The process is initiated at the source node broadcasting *try* messages which contain the set of destination nodes. While these *try* messages travel through the network, cumulative costs are computed by adding the cost of each link passed. Each node receiving a *try* message from a neighbour broadcasts the message to all other neighbouring nodes. A number of variables within the message contain information identifying the followed path and its cumulative costs. The *try* messages travel until they reach a destination or until they are stopped by some timing mechanism. A destination node sends response messages, after it has received *try* messages, by sending *connect* messages via the cheapest path to the source [JIA93].

Complexity

The computational complexity of flooding algorithms is proportional to the maximum length of paths along which messages travel. At path building stage, the longest path from which a destination node may be reached by a *TRY* packet consists of all of the links in the network. Since the maximum number of nodes is $1/2(|N|^2 - |N|)$, this is also the length of such a path. If the path happens to be a viable one, a *CONNECT* packet is generated, and travels along the same route back to the source. Therefore, the worst case performance is $O(|N|^2)$.

As can be concluded, flooding algorithms merely *collect* shortest path information instead of computing appropriate routing trees. The collected information can be used by arbitrary routing algorithms, e.g. the Minimum Destination Cost algorithm.

4.3.5 Performance of non-global algorithms

Results of experiments with MINDC, NH1 and NH2 can be found in [BHA83]. Average network cost (NC) performance of these algorithms is expressed by a comparison between the network cost of the algorithm in consideration and network cost of the MST algorithm. As mentioned earlier, network cost of the MST algorithm is typical within a factor 1.05 from optimal.
The experiments were executed on graphs differing in size: 19 node, 38 node and 100 node graphs. The worst case performance of MINDC, NH1 and NH2 is bounded by 15, 10 and 15 percent worse than MST's, respectively for 19 node graphs. For 38 node graphs, the worst case performance of MINDC, NH1 and NH2 is bounded by 20, 15 and 18 percent worse than MST's, respectively. The network cost performance in case of 100 node graphs decreases with an increasing number of destination nodes for all three algorithms. Their network cost performance becomes more than 30 percent worse than MST's for routings with more than 30 destinations. On average, NH1 performs better than NH2 and NH2 performs better than MINDC (see diagrams in [BHA83]). From the results given in [BHA83] it can be concluded that the NC performance for small numbers of destinations (< 6) is bounded by 15 percent worse than MST's for all three algorithms.

4.4 Extension of MDR algorithms for mutative routing

4.4.1 Mutative algorithms

In chapter three it was concluded that mutative routing is required in order to support adaptive connections for participants joining or leaving established communication sessions. There is an additional motivation for mutative p-mp routing. Generally, telecommunication network deficiencies cause link failures or node failures, whether they occur at physical level or functional level. In case of p-p connections, network failures damage the whole connection, so that a new (complete) connection must be set up. However, in case of p-mp connections, a local network failure possibly damages only a small part of the connection. Evidently, repairing the destroyed part of the connection is a more favourable approach than aborting the whole connection and set up a new one. The latter solution affects all participants of the communication session while the repair approach does not unnecessarily touch the quality of service of remaining connection endpoints.

In the collection of available MDR algorithms described in the previous section, there is only one algorithm suitable for mutative routing, namely the Weighted Greedy algorithm. The Weighted Greedy algorithm is a source node algorithm and requires thus global network information to make routing decisions. All other algorithms perform merely static p-mp routing. However, both the centralized control algorithms and the distributed control algorithms can be modified such that they perform mutative routing as well. The following sections describe extensions of available algorithms making these algorithms suitable for mutative routing.

4.4.2 Proposals for mutative routing with intermediate node algorithms

Modifications of available p-mp intermediate node algorithms must be based on the fact that an existing connection possibly provides a part of the required path from the source to the new destination(s). Therefore, the cost functions of the algorithms must be altered such that only new links and new nodes bring in additional costs.
The intermediate node algorithms can be modified in such a way that 'add' operations need not be constrained to only one new destination at a time. In accordance to the original form, sets of destinations can be added in the same way as a complete new connection is configured.

Nearby Heuristic 1
Assume that an already existing routing R, from source s to destination node set D_e has to be extended with destination set D_a. The same algorithm is used to generate the new part of the routing. However, the computation of *local cost* (part of the total estimated cost) of the new routing is modified. Only links to neighbouring nodes not in R contribute to *local cost.*
In case of equal costs, links and nodes in R should be given a higher priority, preventing unnecessary fan out.
The extended version of NH1 for mutative routing is described in pseudo code as follows:

Nearby Heuristic 1 extended
function **ForwardConnect** ($D_{current}$)
define:

cn = current node
$mc(x, y)$ = cost of shortest path between nodes x and y
$n_1,..n_j$ neighbours of cn except originating node
$D_1,..D_j$ sets of destination nodes to be assigned to $n_1,..n_j$
begin {ForwardConnect}
  if $cn \in D_{current}$ then $D_{current} := D_{current} - cn$
  if $D_{current} \neq \varnothing$ then begin
    consider each possible combination $(n_1, D_1)..(n_j, D_j)$
    compute TotalEstimatedCost =

$$\sum_{1 \le i \le j} e_i\, u_i\, c(cn, n_i) \; + \; \sum_{d_i \in D} mc(n(d_i), d_i)$$

      where e_i is a boolean indicating whether D_i is empty
      and u_i is a boolean indicating whether n_i is already in use
    select the combination with lowest TotalEstimatedCost
    for $1 \le i \le j$ do
      if not e_i then begin
        if not u_i then connect n_i
        Forwardconnect (D_i)
      end
  end {if}
end {ForwardConnect}

Main
begin
  start at source node
  ForwardConnect(D_a)
end.

Figure 22 shows the result of this algorithm used to add two nodes {E, K} to an existing connection {A, F, I, J}.

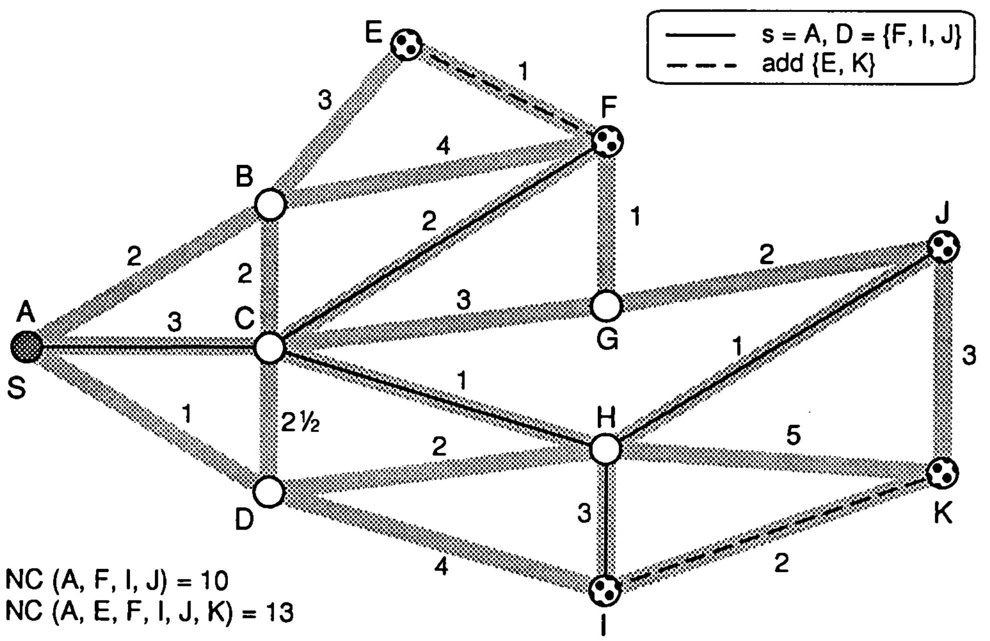


Figure 22 Example of adding nodes with the NH1 algorithm

Nearby Heuristic 2

The modification of this algorithm consists of adding an extra step which first tries to assign the new destinations to neighbouring nodes already in use.
A new group of destinations D_a is added by the existing connection R as follows. Start at the source node. Assume that N_R is the set of already connected neighbouring nodes of source node s. Source node s first tries to assign the destination nodes to N_R. The remaining new destinations are then assigned to the other neighbouring nodes. The analogous procedure is performed at the intermediate nodes.
The extended version of NH2 for mutative routing is described in pseudo code as follows:

Nearby Heuristic 2 extended

```
function ForwardConnect (D_current)
define:
cn = current node
mc(x, y) = cost of shortest path between nodes x and y
D_i set of new destination nodes to be assigned to neighbour n_i ∈ R
begin {ForwardConnect}
  if cn ∈ D_current then D_current := D_current - cn
  for each d ∈ D_current do
  begin
    select n_i' providing lowest mc(n_i', d) ≤ mc(cn, d)
    D_current := D_current - d; D_i := D_i + d
  end
  for each non-empty D_i do ForwardConnect(D_i)
  while D_current ≠ ∅ do
  begin
    select d closest to cn
    connect neighbour n' on shortest to d; D_current := D_current - d
    GroupToForward:= d
    for each remaining d in D_current do
    begin
      if mc(n', d) ≤ mc(cn, d) then add d to GroupToForward
      D_current := D_current - d
    end
    ForwardConnect(GroupToForward)
  end {while}
end {ForwardConnect}

Main
begin
  start at source node
  ForwardConnect(D_a)
end.
```

Figure 23 gives an example of the dynamic use of Nearby Heuristic 2.

Minimum destination cost

Modifications of the minimum destination cost algorithm are constrained by available 'alternate path' information. If only the shortest paths to other network nodes are known, the algorithm simply constructs the routing tree by selecting the neighbour at the shortest path, irrespective eventual links already in use.
Taking advantage of links already in use requires additional information about alternate paths. Using an alternate path can be cheaper if the first link on that path is already part of the

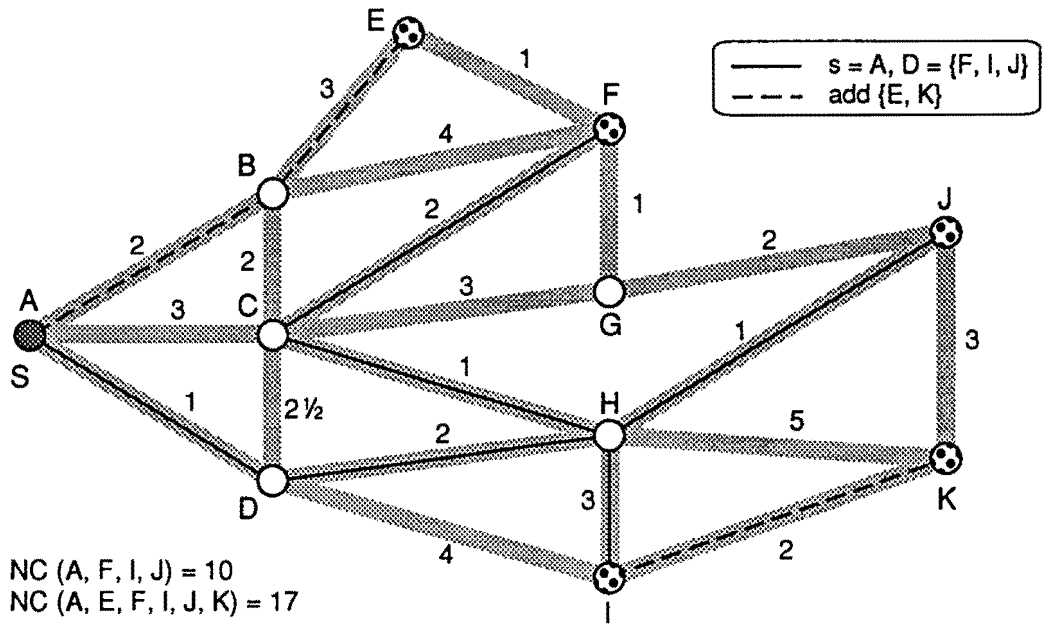


Figure 23 Example of adding nodes with the NH2 algorithm

connection. The computation of the difference in cost requires also the knowledge of the link costs of links attached to the node in consideration.
However, having this additional information, nearby information as required by NH1 and NH2 can be computed. Therefore, a modification of the MINDC algorithm for mutative routing is not sensible.

Loopings

With the above described extensions, it is possible that loopings occur. These can be eliminated by deleting the new path as far as this path was reserved for the *current* new destination nodes.

Deleting nodes

The removal of destination nodes as a result of user departures must be performed such that connections to remaining users stay untouched. This can be done by deleting only the leaf ends of the routing tree connecting the destination nodes to be removed.

4.4.3 Dynamic use of centralized algorithms

The centralized algorithms MST, RS and WONG can also be extended for the computation of additional routes to add new nodes to an existing connection. The idea is straightforward. Construct a new graph in which the nodes and the links already in use form a 'super' node and apply the same algorithm to this graph where the 'super' node operates as source node.

4.5 Implementation improvements

The amount of communication overhead information required by routing algorithms is, among others, determined by the size of the network. The larger the number of nodes and links, the

the more information must be transported and stored in databases. This section investigates how scaling properties of routing algorithms can be improved.

4.5.1 Reduction of communication overhead for non-global algorithms

The previous sections described, among others, some non-global MDR algorithms. The intermediate node algorithms NH1 and NH2 require additional 'nearby' information, which was defined as the shortest path information of neighbouring nodes. The MINDC algorithm requires only shortest path information which was defined as the minimum cost of reaching a particular node and knowledge of the neighbouring node on the shortest path.
Presume that, except primary (shortest) paths, also alternate path information is available to the routing algorithm. In addition, assume that costs of links attached to a given node are available at the node in consideration too. Then nearby information can be computed by simply subtracting the cost of the first link (in primary or alternate paths) of the destination cost.
As a result, less network status information has to be exchanged between network nodes.

4.5.2 Scaling problems

Most routing algorithms intrinsicly suffer from one major drawback: degrading performance with increasing network size [BHA83, DIMI94]. That can be caused by increased communication overhead (centralized control algorithms) or decreased network cost performance (distributed control algorithms).
Centralized control algorithms require global information of the entire network. Therefore, a growing network will increase the communication overhead which is needed to maintain network knowledge at the controlling nodes. This effect is amplified in the case utilization ratio's of the network links make part of the information base in use. Another drawback for centralized routing is the fact that the paths transporting the routing decision information will become longer with larger networks. Longer paths will introduce extra delays regarding the execution of the routing decisions. Furthermore, the computation time at the controlling node will increase according to the time complexity of the algorithm.

Non-global information algorithms are less sensitive to increased communication overhead because they simply require less network information compared with global information algorithms. However, non-global routing does require the network information to be available in each network node, while global routing needs only a few control nodes to have the information.
Non-global information algorithms have a disadvantage regarding their network cost performance. This bad scaling property is caused by the fact that they have only local information. This can become a serious drawback in the case of large networks, see [BHA83].
Especially in case of routings with many destinations (≥10) the network cost performance deteriorates severely.

4.5.3 The multidomain concept

The above described scaling problems can be unravelled by dividing the telecommunication network in a number of functional domains, also called partitioning. A functional domain is an independently operating subnetwork; several subnetworks construct a complete network. Except the scaling problems, there is an additional motivation for regarding the network as a *multi-domain network*. Telecommunication networks in practice can namely be non-homogeneous for several reasons:

- the network is owned and managed by more than one operator
- the network contains subnetworks built on distinct technologies (e.g. old and new)

Interworking within such a non-homogeneous network environment clearly requires a multi-domain concept.

The multi-domain concept offers a potential solution to the problems of increasing network size for routing systems and the problem of coping with subnetworks in an heterogeneous telecommunication environment [DEER88, DIMI94, EST93]. In a multi-domain network, each domain is controlled by its own *Network Control Centre* (NCC). The NCC is a functional control entity, the exact implementation in the network is not important here. Centralized routing control is an example of a possible control function for the NCC. Management and control activities of the NCC include cooperation among other NCC's. This distributed approach involves peer-to-peer coordination of the various NCC's involved, while an hierarchical approach introduces an *Integrated Network Control Centre* (INCC). This INCC supervises NCC's and their interaction. Several individual controlled domains are interconnected via so called *gateway* links.

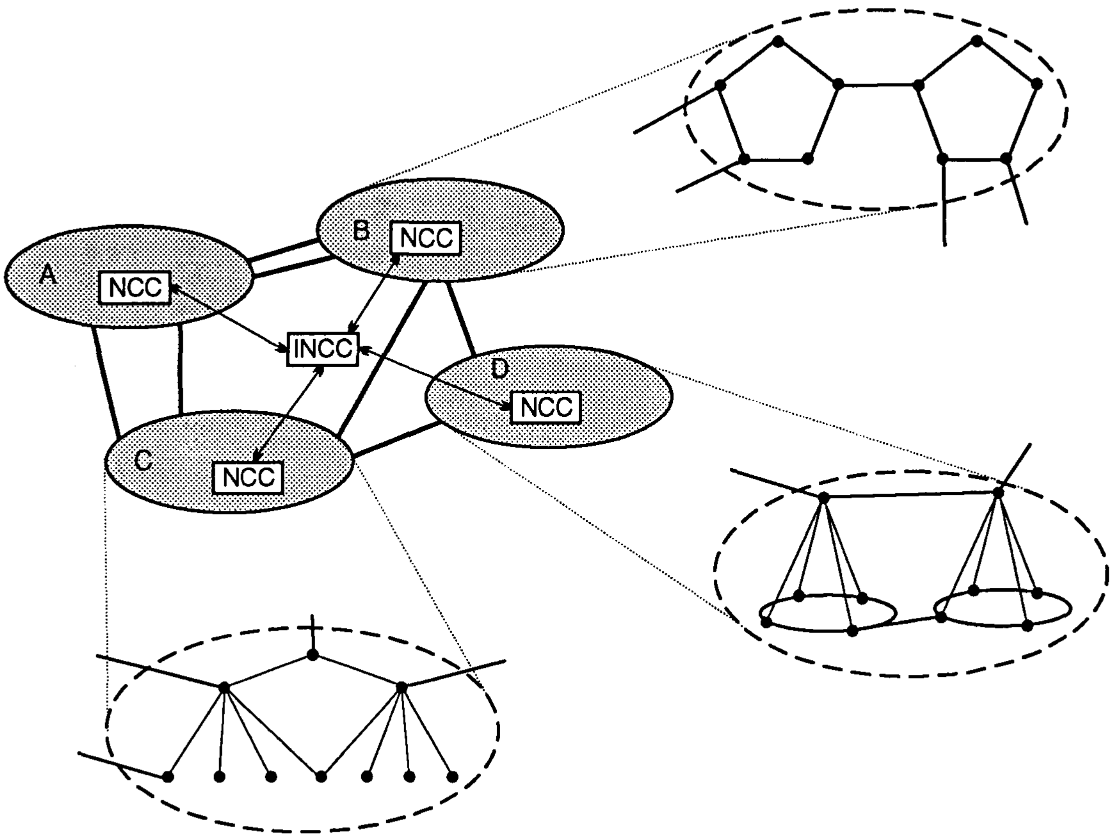


Figure 24 The multidomain concept

Figure 24 illustrates the multidomain concept with a two layer hierarchical structure. The network consists of four subnetworks A, B, C, D that interwork either peer-to-peer or in a client-server manner. In the latter case, the INCC is the server that provides for example an inter domain connection requested by one of the NCC's, who act as clients.
The hierarchical structure can be extended to several layers but can also reduced to one layer. The latter is the case when no supervising entity is available, e.g. when national networks are coupled on basis of equality, and thus constructing an international network without a supervisor.

Connection management
Connection management in multi-domain networks requires a system of interworking between controlling entities of the subnetworks such as provided by the Telecommunication Information Networking Architecture [TINA93]. Connection management for intra domain connections is performed within one subnetwork. Connection management for inter domain connections requires additional communication between subnetworks (peer-to-peer or client-server). Individually controlled subnetworks use their own routing algorithms for set up of intra domain connections independent of other subnetworks. Inter domain connections require higher level routing algorithms. Peer-to-peer communication for inter domain connections requires a distributed approach, while the client-server method enables use of centralized control.

The advantage of network partitioning for routing algorithms is clear. Global information algorithms require less communication overhead and storage capacity. The advantage for non-global information algorithms is that the worst case network cost performance is bounded by the limited size of a subnetwork.

4.6 Comparison of MDR algorithms

4.6.1 Overview

The multiple destination algorithms described in this chapter are summarized in table 4.2. Algorithms NH1e and MH2e are extended versions of NH1 and NH2.
Network cost (NC) performance indicators are not included in this table. The centralized control algorithms (MST, RS and DAA) have near optimal NC performance, WGA's NC performance is typical within a factor 1.5 from optimal and the NC performance of the non-global algorithms depends on the size of the network and the number of destinations. On average, NH1(e) has the best performance of the non-global algorithms followed by NH2(e).

Table 4.2 Overview of MDR algorithms and their main properties

MDR algorithms	Network information	Mutative	Appropriate for directed graphs	Reference
MST	global	possible	no	[MEHL88]
RS	global	possible	no	[RAY86]
WGA	global	yes	yes	[DOAR93]
DAA	global	possible	yes	[WONG84]
MINDC	local	no	yes	[BHA83]
NH1	nearby	no	yes	[BHA83]
NH1e	nearby	yes	yes	section 4.4
NH2	nearby	no	yes	[BHA83]
NH2e	nearby	yes	yes	section 4.4
(Flooding)	local	difficult	yes	[JIA93]

The flooding algorithm is placed between brackets because it composes merely shortest paths between single nodes and does not build routing trees.

4.6.2 Judgement of algorithms

As can be seen in the previous, centralized control algorithms have the best network cost performance and should therefore be favoured where possible. However, not in all cases centralized control algorithms can be used.
In chapter 4 several routing strategies were distinguished in order to meet different requirements of services with distinct characteristics. State-dependent routing was defined in order to route services with short holding times such that network resources are used optimally

while lowering call blocking probabilities. State-dependent routing requires for this task momentary (near real time) network status information (e.g. occupancy ratio's). Time-dependent routing requires only long term network status information. State-dependent routing using centralized algorithms places a burden on communication overhead caused by near real time network status updates from the whole network. Non-global intermediate node algorithms require only updates for their shortest path information tables and not of the entire network. As a result, centralized algorithms are less appropriate for state-dependent routing than non-global routing algorithms.
The conclusion is that non-global information algorithms are suitable for state-dependent routing. Using centralized control algorithms for state-dependent routing requires additional network constraints. As concluded in the previous section, a multi-domain network enables detailed partitioning so that individual subnetworks become small components. In a subnetwork limited in size, real time communication overhead for centralized routing algorithms can also be kept limited. Therefore, centralized algorithms can be used for state-dependent routing in multi-domain networks under the constraint that the subnetworks are limited in size.

Time-dependent routing is required to adapt the average network load to recurrent utilization patterns, as explained in the previous chapter. In contrast to real-time network status information, the long term nature of status information for time-dependent routing enables usage of centralized databases without introducing excessive communication overhead. These databases for network routing information can then be used by centralized routing algorithms. Therefore, centralized algorithms are suitable for time-dependent (and state-independent) routing and should also be used for time-dependent routing because they have the best network cost performance. For example, semi-permanent broadcast VP's in an ATM trunk network should be routed state-independent, time-dependent by a centralized routing control entity.

Which algorithm (distributed control)?

Consider the situation that the network cannot be partitioned caused by the lack of an information network architecture that provides communication between subnetworks. In this case, a non-global MDR algorithm should be selected for state-dependent routing. Because alternate path information must be available for state-dependent routing, nearby information can be computed as described in section 4.5.1. Therefore, NH1 or NH2 are preferred above MINDC because they have better network cost performance.
The computational (time) complexity of NH1 increases exponentially with the number of neighbouring nodes. The computational complexity of NH2 increases only linearly with the number of neighbouring nodes. It depends on the processor speed of the computers on which the algorithms run whether this distinction between both algorithms brings NH2 in favour. If the computation time does not contribute significantly to the overall connection set up delay, then NH1 should be favoured due to its better NC performance.

Which algorithm (centralized control)?

Point-multipoint services cause imbalanced network loads. Therefore, the network must be modeled as a directed graph. Thus, the Minimum Steiner Tree and the Rayward Smith algorithms are not useful for centralized p-mp routing. So the Dual Ascent Algorithm and the Weighted Greedy Algorithm remain. The Dual Ascent Algorithm has a near optimal NC performance while the Weighted Greedy Algorithm performs only within a factor 1.5 from optimal. Therefore, the Dual Ascent Algorithm should be selected for centralized computing of p-mp connections.

4.7 Conclusions

Efficient network resource utilization for p-mp services requires computation of optimal (shortest) multiple destination routes through the network. Determination of such an optimal routing tree is known in the mathematical branch of graph theory as the Minimum Steiner Tree problem, which belongs to the class of NP-complete problems. A number of heuristic algorithms are available from the literature, which generate solutions for the MDR problem based on a graph model of the telecommunication network. Because point-multipoint connections load the network in an asymmetrical way, the network must be modeled as directed graph in order to be able to make distinction between both directions of network links and nodes.

Routing algorithms can be classified as either centralized control algorithms or distributed control algorithms. Centralized control algorithms require knowledge of the entire network topology, indicated with global information, while distributed control algorithm need only knowledge of a part of the network, indicated with local information.

Multiple destination routing algorithms possibly suffer from two scaling problems:

- centralized control algorithms require large databases and cause excessive communication overhead in large networks
- distributed control algorithms have bad network cost performance in large networks

A possible solution for these two scaling problems is network partitioning, which creates a multi-domain network. A multi-domain network is a hierarchical structured collection of (loosely) coupled subnetworks. Each subnetwork controls its own domain and uses routing algorithms independently of neighbouring subnetworks. The management of multi-domain networks requires a management information architecture like TINA.

A basic requirement on MDR algorithms is that they must be able to perform in mutative routing strategies. There is only one centralized control MDR algorithm which satisfies this requirement and which has a relatively bad network cost performance compared with other centralized control algorithms. Therefore, the available distributed control algorithms were extended for use in mutative routing. Also a solution is given for mutative routing with the remaining centralized control algorithms.

Most centralized control algorithms generate near optimal routing trees, thanks to their overall network knowledge. Therefore, centralized control algorithms should be used where possible. However, centralized control algorithms are not suitable for state-dependent routing strategies unless the network is partitioned to a sufficient level of detail. In any case, centralized control algorithms should be used for state-independent routing. The Dual Ascent algorithm is the best choice because it is the only algorithm which models the network as a directed graph.

With respect to least cost routing trees, distributed control algorithms perform worse than centralized control algorithms. However, due to their limited requirements on available network status information, distributed control algorithms should be preferred for state-dependent routing of connections.

5 ATM connection management

Connection management plays an important role in network resource management because routing procedures select network resources to be allocated for requested connections. This chapter describes how connection management can perform resource management tasks using multipoint routing algorithms for both state-dependent and time-dependent routing strategies. The main objective is to integrate routing procedures for all types of connections (p-p, p-mp and mp-mp) for all types of traffic in one congestion preventive routing approach.

5.1 Link resource management

The wide range of heterogeneous services to be supported by ATM networks consists mainly of services requiring quality of service guarantees. Many telecommunication service applications will depend on the real-time transport of information through the network. This makes network resource reservation imperative. Realization of connections through the network requires reservation of link bandwidth, buffer capacities and node processing capacities. Controlling access to these network resources is particularly important for achieving quality of service objectives and grade of service objectives, see section 2.1.1. The latter relates to call blocking probabilities for the offered traffic. Resource management controls are required to control access to network resources. They can be implemented at a number of levels in an ATM network and can be applied to Virtual Path and Virtual Channel connections, as well as individual cells within a particular connection. Resource management controls at VP and VC level will be analyzed in the following.

Whereas the initial idea underlying the development of ATM has been that of a complete statistical multiplexing, control architectures oriented towards a more 'structured' resource allocation have also been considered, where portions of a resource (e.g. bandwidth) can be temporarily and dynamically assigned to traffic with specific characteristics or requirements [BOL94, GUP94]. In ATM, traffic with specific characteristics can be advantageously grouped in so called Quality of Service classes, see also [I211]. The VP concept provides namely the possibility to combine VC's belonging to a particular QoS class in one VP connection. A Virtual Path must satisfy the QoS demands of the VC with the highest QoS requirements carried within that VP. Therefore, it is preferable to keep differences in QoS requirements of individual VC's within one VP as low as possible. This means that several VP's are required in order to support efficiently distinct QoS classes.

Example

Some applications, such as interactive voice and video applications, are delay sensitive, but not loss sensitive. Others, such as file transfer and electronic mail, are not delay sensitive, but loss sensitive. The Virtual Path concept can be used to separate traffic with such different QoS requirements.

When several VP's share resources on a transmission link, some resource control mechanism is required which allocates available resources to the VP's [BOL94]. Figure 25 illustrates a

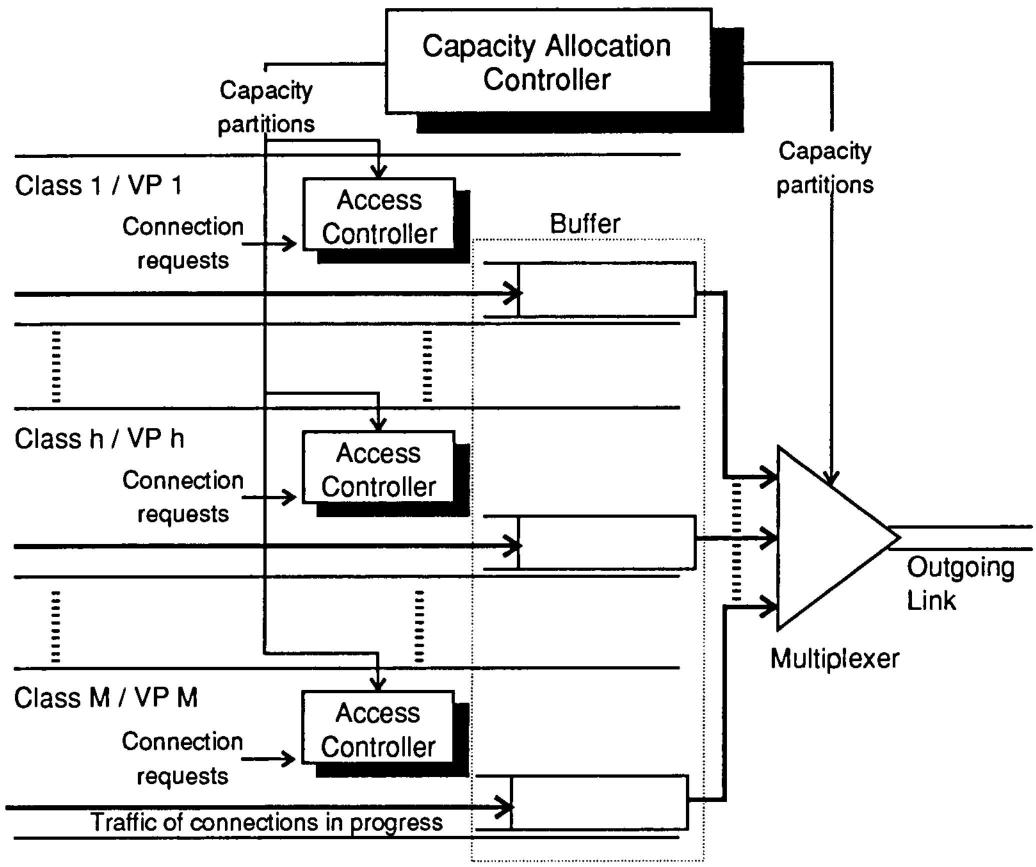


Figure 25 The overall control structure of a link inside an ATM node

possible overall resource control structure inside an ATM node. Available resources are partitioned and allocated to the VP's carried by an outgoing physical link. The capacity allocation controller creates these partitions according to the resource needs in distinct QoS classes and assigns these resources to the corresponding VP's.

Due to statistical traffic fluctuations, the resource needs per VP will be dynamic. These traffic fluctuations are caused by fluctuations in the number of VC's contained in one VP and traffic burstiness in case of variable bit rate VC's. Therefore, the capacity allocation controller must also perform in a dynamic way. Once the resources per VP are allocated, access controllers are required for each VP to shield the QoS of connections in progress. These access controllers decide whether new connection requests can be complied with.

Consider the situation where a particular VP is congested and that other VP's using the same link are not fully occupied yet. Two mechanisms can be used now in order to save additional connection requests:

- modify VP capacity partitions such that the congested VP gets more resources
- allocate resources for the requested VC's in a VP used for another QoS class which is not yet fully occupied

Modification of VP capacity partitions puts additional dynamic requirements on the capacity allocation controller. However, the advantage above the second solution is that traffic of services with distinct characteristics remains separated conform the basic objectives for VP use. The advantage of the second solution is that it can easily be implemented in state-dependent routing procedures. Virtual Paths belonging to other QoS classes form then additional alternate paths for the routing algorithm in operation. Because all routing algorithms minimize a particular cost function, this cost function can be used to prefer VP's with a high degree of correspondence between the requested QoS and the delivered QoS. Such a VP selection criteria can be integrated in routing procedures by adding a cost parameter in the cost function for the degree of QoS correspondence.

The conclusion is here that state-dependent VC routing procedures can contribute significantly to the overall resource control of individual network links. Because VC resource control performs complementary to VP resource control, state-dependent routing procedures lower dynamic requirements on VP resource control.

5.2 Integrated resource allocation mechanism

As concluded above, routing procedures can play a major role in overall network resource control. On the routings determined by routing procedures, available resources must be allocated in order to transport user and network information. The amount of network resources to be allocated depends on whether the information flow is either unidirectional or bidirectional and the required QoS level. Required network resources for distinct services vary significantly, mainly depending on the traffic type, e.g. video (several Mbit/s) or speech (several Kbit/s). The following presents an integrated approach for routing reserved network resources for all types of connections.

Chapter 3 distinguished state-dependent and time-dependent routing strategies based on the fact that traffic fluctuations occur in short term and in long term intervals. The previous chapter described multiple destination routing algorithms and their applicability in both state-dependent and time-dependent routing strategies. The amount of network status information required by these algorithms was a main criteria for classification of the algorithms. The following analyzes more specifically the required contents of network status information on which routing decisions should be based such that an integrated approach as described above can be achieved.

Routing algorithms make decisions using a particular cost function. This cost function computes the relative cost of using specific (functional) links for a requested connection. This cost function provides now the possibility to make distinction between several QoS levels and distinct symmetries in case of bidirectional information flow.
Regarding the requested QoS, the following parameters should be included in the cost function, according to the subattributes of the QoS attribute defined in section 2.1.1:

- peak rate
- average rate

- burstiness
- peak duration

In case of bidirectional information flows, both directions must be considered in the cost function, while in case of unidirectional information flow only one direction has to be considered. When network resource demands in both directions are fixed for the entire connection, both the centralized and the intermediate node routing algorithms are able to model the situation by including the QoS parameters for both directions in the cost function. Resource demands in both directions are fixed for p-p connections and unidirectional p-mp connections. However, in case of bidirectional multipoint-multipoint connections, resource demands are not fixed for the entire direction. Multipoint-multipoint connections can be used for conference services as described in section 2.2.2. In mp-mp connections, the downstream information flow direction requires the same resources for the entire connection, but in the upstream direction the required network resources accumulate caused by concentration of the multipoint-point traffic when it travels to the root of the connection.

Intermediate node MDR algorithms, such as NH1 and NH2, have inherently the property of handling varying resource demands caused by concentration of upstream information traffic. The reason is that the algorithm in operation simply knows the number of destinations to be connected by the 'current' decision making node in the routing tree. Therefore, also the required upstream network resources can be computed based on the (local) number of destinations to be connected.

Without modifications, centralized MDR algorithms, such as DAA and RS algorithm, cannot cope with the problem of varying reverse directed traffic flows in multipoint-multipoint connections. This problem was addressed in [VERM93], where some modifications to the Minimum Steiner Tree heuristic are presented in order to construct efficient routing trees which include reverse information flows.
However, under the following assumptions, the problem of reverse directed traffic flows in routing trees can be simplified such that mp-mp connections can be routed in the same manner as unidirectional p-mp connections. Assume that state-dependent routing mechanisms are in operation in order to improve network resource utilization. Furthermore, assume that balanced link utilization is achieved by the proposed extension of the cost function: links with imbalanced traffic load are preferred by unidirectional connections such that link load balance is improved. Large numbers of connections will then guarantee that network link loads are actually in balance. Assume also that, on the average, in the upstream direction less network resources are required than in the downstream direction. Under these assumptions, the mp-mp routing problem can be modeled as a unidirectional p-mp routing problem. The motivation is that the resulting routing tree will then be optimized to the most demanding (downstream) resource requirements. The balanced link loads guarantee that in the opposite direction enough resources are available for the upstream traffic flows which require less resources.

The conclusion is that the routing of both unidirectional and bidirectional point-multipoint connections can be performed by intermediate node and centralized control routing algorithms. A prerequisite is the use of a bidirectional cost function which must also be able to support several quality of service levels. Therefore, a number of QoS parameters must be included in this function.

In order to compute efficient routes, routing algorithms require network link costs in case of centralized algorithms and accumulated link costs of predetermined paths in case of intermediate node algorithms. For providing this required network status information, it is of major importance to make distinction between state-dependent and time-dependent routing strategies. Time-dependent routing procedures are applicable to both Virtual Path and Virtual Channel resource control and use primary path information. State-dependent routing procedures use alternate path information; they are applicable to routing problems of short-lived connections at VC level. The (semi)-permanent character of Virtual Path connections is the reason that state-dependent routing is not sensible for VP's.

Primary path information

Primary paths are prearranged shortest paths irrespective of momentary network status. Primary paths should be computed with time-dependent and state-independent routing algorithms. Provided that the network is virtually subdivided using distinct VP's for distinct QoS classes, each QoS class requires its own primary paths. Momentary occupancy degrees of network resources are not needed for construction of primary paths. Time-dependent construction of

primary paths requires average periodical network status information regarding:

- bidirectional occupancy degrees of available buffer and transmission capacities
- relative proportions of resource usage by distinct QoS classes

Primary paths for distributed control algorithms can be constructed using centralized routing algorithms.

Alternate path information

Alternate paths are not predetermined like primary paths, but alter continuously with altering network resource occupation degrees. Network status updates are required with short time intervals to provide the routing algorithm the needed information. It was already argued that using centralized control algorithms for alternate routing burdens the network with huge resource control communication overhead (except in the case of detailed network partitioning using the multidomain concept). Distributed control algorithms require prearranged paths which must be reconfigured within the permitted time interval. The update mechanism for distributed control algorithms can be implemented using flooding algorithms (see Appendix E). When the network is lightly loaded, the shortest alternate paths resemble primary paths.

5.3 Preventive congestion control based routing

The value of state-dependent routing over state-independent routing for calls with short holding times was already made obvious in chapter three. When there are insufficient resources available on a call's primary path, allowing a call to complete on an alternate path prevents it from being lost.

Less obvious are the problems that uncontrolled state-dependent routing can cause in a network. Careful studies of state-dependent routing even under symmetric scenarios show that uncontrolled state-dependent routing can actually do much worse than state-independent routing when the load on the network is beyond a certain critical load, see [SIB94]. The exact value of this critical load depends on several factors: alternate path lengths, network size, graph structure of the network, etc. The key to understanding why state-depending routing can have a deleterious effect is this: typically, an alternate-routed call will use more resources than a call routed on the primary path. Acceptance of an alternate-routed call can cause other calls to be blocked on their primary path and force them to a less efficient alternate path. This in turn can aggravate the problem of finding primary paths for other calls, resulting in an even larger fraction of calls choosing alternate routes. Such an avalanche effect drives the network into a high-blocking operating region. This behaviour is not significant when the network is lightly loaded, since the probability of blocking on the primary path is small enough that the fraction of calls using alternate routes remains small. However, under heavy loads the network can be driven into an inefficient state, as described in [BAHK94, SIB94].

Since uncontrolled state-dependent routing can lead to undesirable behaviour at high loads, a mechanism is required which prevents such behaviour. A well-tested mechanism, known as *trunk reservation* or *state-protection*, provides the desired robustness. It is a method by which primary traffic is given priority over alternate traffic. Thus, state-dependent routed calls experience a higher blocking probability than state-independent routed calls. Experiments show that this scheme is robust, in that a state-protection level optimized for a specific loading situation works well under variations in load [SIB94].

It was already concluded in chapter three that both state-dependent and state-independent routing mechanisms are needed at the same time. The reason is that calls with long holding times or high network resource requirements should be routed state-independent in any case. Routing algorithms can easily switch from state-dependent to state-independent routing by using the corresponding primary or alternate path information tables. The selection of one of both mechanisms can be done simply by introducing a *flag* in the call request message which indicates which routing paths should be used. This flag provides now the possibility to give certain calls higher priority in case of network congestion. That can be useful in the following cases:

- set up of subscription based connections
- set up of high revenue calls
- set up of calls that were already blocked in a previous attempt.

5.4 Conclusions

Efficient use of the ATM Virtual Path concept necessitates implementation of distinct Virtual Paths for several QoS classes containing distinct traffic types such as speech, video and data. Functional separation of the network using Virtual Paths causes additional requirements on storage and communication capacities for routing information used by routing algorithms.
Link resources are allocated to different VPs using a capacity allocation mechanism. In case of congested VPs, two methods are available to accept a new requested VC connection. The first is a dynamic capacity allocation scheme and the second is state-dependent routing of VCs, where other VPs form alternate paths. The advantage the former method is that distinct traffic types remain separated. The advantage the latter method is that it lowers the dynamic requirements on the capacity allocation mechanism.
Routing algorithms minimize a particular cost function. Centralized algorithms use costs per link while distributed control algorithms use accumulated costs of paths. The cost function provides the possibility to integrate the routing of p-p, p-mp and mp-mp connections with different QoS levels within one approach. For this task, the desired QoS parameters for both directions of the requested connection are required just as bidirectional (accumulated) link costs. In case of state-dependent routing, these link costs form primary paths. In case of state-dependent routing, the link costs form alternate paths and must be updated frequently in time.

Multipoint-multipoint connections put additional requirements on routing algorithms caused by their non-fixed upstream capacity requirements. Distributed control algorithms have inherently the property of handling these varying demands. Under a number of (practical) assumptions, centralized control algorithm can also route mp-mp connections without modifications. Preventive congestion control based routing requires implementation of both state-dependent and time-dependent routing strategies. Uncontrolled state-dependent routing leads to instabilities when network congestion occurs. When the network load reaches a specific level, no state-dependent routing should be allowed to prevent an avalanche effect caused by less efficient resource utilization of state-dependent routed connections, as concluded in [BAHK04, SIB94]. The required mechanism, called state-protection, can be implemented in multipoint routing algorithms by switching to primary path information in congestion situations. This mechanism also provides the possibility of prioritization of connection requests.

6 Conclusions and recommendations

Conclusions

New transmission and switching techniques, like ATM, support the integration of a wide range of new multipoint services in the public network. The required multipoint connection management functionality must meet service specific requirements and general network management objectives. An extensive literature study made obvious that **multipoint routing**, which is essential for connection management, causes problems when both categories of requirements must be met at the same time.

Service specific multipoint connection requirements are structured by the separation of two classes of multipoint services. The first class consists of **interactive conference** services with the following connection requirements.

(a) The number of connection endpoints is limited.

(b) Bidirectional asymmetric multipoint-multipoint connections or unidirectional point-multipoint connections are needed.

(c) Complete multipoint connections must be set up on demand or based on subscription contracts.

(d) A multipoint connections in progress must be extended and reduced on demand caused by user join and leave events respectively.

The class of **distributive services** has distinct connection requirements.

(a) The possible numbers of connection endpoints varies between a few and 'unlimited'.

(b) The connection is of the type unidirectional point-multipoint.

(c) Parts of the multipoint connection are semi-permanent.

(d) Dynamic requirements depend on the implementation in the access network.

Required transport capacities are determined by the traffic type and vary from a few kbit/s for speech or information retrieval services to a few Mbit/s for good quality video services.

Basic network resource objectives must be met because connection management plays a major role in network resource management. Basic objectives of network resource management in general are optimal network resource utilization, prevention of congestion situations, maintaining quality of service guarantees for connections in progress and keeping control overhead low. From existing telephony networks it is known that the routing of connections can contribute significantly to these general objectives. Therefore, routing strategies need to be developed which adapt to traffic fluctuations in integrated services networks. Traffic fluctuations occur at short term (reflected by actual occupancy degrees) and long term intervals (reflected by average, periodical patterns).

The variety in multipoint service characteristics necessitates two distinct routing strategies in order to adapt the routings of all types of multipoint services efficiently to these traffic fluctuations. **State-dependent routing** is required for adaption to momental fluctuations and **time-dependent routing** is required for adaption to long term traffic fluctuations. Asymmetrical multipoint connections cause unbalanced link loads. Therefore, multipoint routing procedures must perform a **link balancing** mechanism. This can be achieved by including additional parameters for both link directions in network load indicators which are to be used for adaptive routing strategies.

Efficient network resource utilization for multipoint connections requires computation of optimal (shortest) multiple destination routes through the network. The determination of such an optimal routing tree is an NP-complete problem. Therefore, heuristic **multipoint routing algorithms** are required in order to compute efficient routings within a polynomial bounded time.
A number of routing algorithms are available, which are either controlled centrally or controlled distributively. Centralized control algorithms generate near optimal routing trees, thanks to their overall network knowledge. With respect to least cost routing trees, distributed control algorithms perform worse than centralized control algorithms, caused by the lack of complete network knowledge. This effect escalates in case of large numbers of destinations in large networks.
Centralized control algorithms can be implemented in a distributed way by creating a **multi-domain network**. A multi-domain network is the result of network partitioning, which divides the network in individually controlled sub-networks. A multi-domain network requires an additional form of communication between the individual sub-networks. This can be done hierarchically or based on equality using a peer-to-peer communication method. Network partitioning and both forms of interaction between sub-networks are supported by TINA. The network cost performance of centrally controlled multipoint routing algorithms implemented in such a distributed system is still an open issue.

Recommendations

Due to the variety in multipoint service characteristics, multipoint routing algorithms are needed for both state-dependent and time-dependent routing strategies. The choice of the routing strategy should be based on session holding times, transport capacity, possible subscription contracts and maximal permitted transport delay. These service characteristics determine the cost of using alternate routes in terms of efficient network resource usage.

Broadcast-like distributive services with large numbers of users can be advantageously supported by special broadcast multipoint Virtual Paths in ATM networks. **Distribution Virtual Paths** in the trunk network separate permanent distributive information streams from dynamic interactive traffic, which eases connection management and enlarges reliability of transport of distributive traffic. User group dynamics have only consequences for the access networks in this case. Distribution Virtual Paths should be routed time-dependently.
Conference services and limited user group multicast services are characterized by a limited number of participants and require therefore a distinct routing approach at Virtual Channel level. The variety in service characteristics of those services makes the implementation of state-dependent and time-dependent routing strategies for the set up of multipoint Virtual Channels imperative. Dynamics in user groups of these services apply to the entire connection. Therefore **mutative** multipoint routing algorithms should be used.

It is recommended that **centralized control** multipoint routing algorithms should be used where possible on account of their near optimal network cost performance. However, centralized control algorithms induce excessive communication overhead in state-dependent routing strategies caused by near real-time network status updates. Therefore, the network must be partitioned in a multidomain network in order to keep the control overhead limited. When no multi-domain management information architecture is available, **distributed control** algorithms become preferable for state-dependent multipoint routing because they need only limited network information for making routing decisions.
The routing of **non-bidirectional** multipoint connections necessitates the use of a directed graph as a model of the network. So the Dual Ascent algorithm is an appropriate centralized control routing algorithm. In case of distributed control algorithms, it is sufficient to extend the network cost function to a bidirectional cost function.
The mutative routing requirements can be met with the **proposed extensions** of both the distributed and centralized control algorithms.

An integrated routing approach is recommended for all types of multipoint services by implementation of bidirectional **quality of service parameters** and bidirectional Virtual Path link costs in the network cost function. In case of state-dependent routing, the bidirectional Virtual Path link costs should include occupation degrees of both link directions, which must be updated frequently. Time-dependent routing requires availability of recurrent patterns of average values of Virtual Path link occupation degrees.
State-dependent routing ameliorates dynamic Virtual Path capacity allocation mechanisms when distinct Virtual Paths contained by one physical link serve as alternate paths. Therefore, state-dependent routing procedures should play a role in the design of Virtual Path capacity allocation mechanisms.
Time-dependent routing is based on the use of primary paths. State-dependent routing uses also alternate paths. Since uncontrolled state-dependent routing leads to network instabilities in case of congestion, preventive congestion control based routing requires a **state-protection** mechanism. This state-protection mechanism switches from alternate path to primary path routing when the network load reaches a specific critical level. Therefore, state-dependent multipoint routing algorithms must also be able to route time-dependently. This is achieved by including primary path information in their network information databases.

Appendix A Literature

[AMMA93] Ammar M.H. et al.
Routing multipoint connections using virtual paths in an ATM network
Proc. of Eleventh annual joint conf. of the IEEE Computer and communications Societies, INFOCOM '93: NETWORKING: FOUNDATION FOR THE FUTURE, San Francisco, March 30 - April 1, 1993, Technical Program Chair: K.S. Vastola, p. 98-105, Vol. 1

[ASH90] Ash, G.R.
Design and control of networks with dynamic nonhierarchical routing
IEEE Communications Magazine, October 1990, p. 34-40

[ASH92] Ash, G.R. et al.
Real-time network routing in the AT&T network, improved service quality at lower cost
Proc. of IEEE Global Telecommunications Conference (Orlando, December 1992), p. 802-813

[BAHK94] Bahk S. and Zarki, M.E.
Preventive congestion control based routing in ATM networks
Proc. of IEEE ICC 1994, p. 1592-1599

[BOL94] Bolla R. et al.
A distributed routing and access control scheme for ATM networks
Proc. of IEEE ICC 1994, p. 44-50

[BREE92] Brandt, D. et al.
In breed verband; wegwijs in nieuwe technieken voor breedbandtelecommunicatie
PTT Research

[CHANG94] Chang Y.H. et al.
An open-systems approach to video on demand
IEEE Communications Magazine, May 1994

[BHA83] Bharath-Kumar K. and J.F. Jaffe
Routing to multiple destinations in computer networks
IEEE Transaction on communications, Vol. COM-31, No. 3, March 1983

[CCIR91] Bergman, S. et al.
The SR report on the MPEG/Audio subjective listening test
CCIR document JIWP-CMTT/1-45
Stockholm, June 1991

[CHEN93] Chen, C.-T., et al.

Hybrid extended MPEG video coding algorithm
Signal Processing: Image Communication, Vol 5, Nos. 1-2, February 1993

[CHOW90] Chow C.H.
On multicast path finding algorrithms
Proc. of 8th annual joint conf. of the IEEE Computer and communications Societies, INFOCOM '90, p. 1274-1283, 1990, Vol. 2

[COST93] COST 211 ter, ATM compendium, Simulation Subgroup
Coding procedures
May 1993

[CSELT95] ETSI NA 5 OAM contribution
Maintenance of Point-to-Multipoint Connections, Proposal for an appendix to Rec I.610
Source: CSELT/STET - Italy, January 1995

[DEE90] Deering S.E. and D.R. Cheriton
Multicast routing in datagram internetworks and extended LANs
ACM Transactions on Computer Systems, Vol. 8, No. 2, May 1990, p. 85-110

[DEE88] Deering S.E.
Multicast routing in internetworks and extended LANs
Proc. of ACM SIGCOMM Symposium August, 1988

[DIMI94] Dimitrijevic D.D. et al.
Routing in multidomain networks
IEEE/ACM Transactions on networking
Vol. 2, No 3, June 1994

[DOAR93] Doar M. and I. Leslie
How bad is naïve multicasting?
Proc. of Eleventh annual joint conf. of the IEEE Computer and communications Societies, INFOCOM '93: NETWORKING: FOUNDATION FOR THE FUTURE, San Francisco, March 30 - April 1, 1993, Technical Program Chairman: K.S. Vastola, p. 82-89, Vol. 1

[DOS93] Doshi, B.T.
Deterministic rule-based traffic descriptors for broadband ISDN: Worst case behavior and connection admission control
Proc. of IEEE GLOBECOM 1993, p. 1759-1764

[EST93] Estrin, D et al.
A protocol for route establishment and packet forwarding across multidomain internets
IEEE/ACM Transaction on Networking, Vol. 1, No. 1, February 1993

[FLOR91] Floren R.
A note on "A faster approximation algorithm for the Steiner tree problem in graphs"

Information Processing Letters, Vol. 38, p. 177-178
Elseviers Science Publishers B.V., North Holland, 1991

[GUP94] Gupta, S. and Gandhi P.P.
Dynamic routing in multi-class non-hierarchical networks
IEEE International Conference on Communications
1994, p. 1390-1393

[HAC94] Hac A.
Distributed multicasting algorithm with congestion control and message routing
Journal on Systems Software, No. 24, p. 49-65
Elseviers Science Inc., New York, 1994

[H261] ITU-T Recommendation H.261
Video codec for audiovisual services at p x 64 kbit/s, December 1990

[HWA92] Hwang F.K and C.S. Richards
Steiner tree problems
Networks, Vol. 22, p. 55-89
John Wiley & Sons, Inc., 1992

[I140] ITU-T Recommendation I.140
Attribute technique for the charterization of telecommunication services supported by an ISDN and network capabilities of an ISDN
March 1993

[I211] ITU-T Recommendation I.211,
B-ISDN service aspects
March 1993

[I311] ITU-T Recommendation I.311
B-ISDN general network aspects
March 1993

[I610] ITU-T Recommendation I.610
B-ISDN operation and maintenance principles and functions
March 1993

[IMA91] Imase M. and B.M. Waxman
Dynamic Steiner tree problem
SIAM Journal on Descrete Mathematics, Vol. 4, No. 3, p. 369-384, 1991

[JAF85] Jaffe J.
Distributed Multi-Destination Routing: The Constraints of Local Information
SIAM Journal on Computing, Vol. 14, No. 4, p. 875-888, Nov. 1985

[JIA91] Jiang X.
Path finding algorithms for broadband multicast

High speed networking, Vol. III, p. 153-164
Elseviers Science Publishers B.V., North Holland, 1991

[JIA92] Jiang X.
Routing broadband multicast streams
Computer Communications, Vol. 15, No. 1
Butterworth-Heinemann LTD, 1992

[JIA93] Jiang X.
Distributed path finding algorithm for stream multicast
Computer Communications, Vol. 16, No. 12, p. 767-775, December 1993

[KAW93] Kawashima, M. et al.
Adaptation of the MPEG Video-Coding algorithm to network applications
IEEE Transaction on circuits and systems for video technology, Vol.3, No. 4, August 1993

[KEY90] Key, B.P. and Cope, G.A.
Distributed dynamic routing schemes
IEEE Communications Magazine, October 1990, p. 54-69

[KIN93] Kinoshita, T. et al.
Variable-Bit-Rate HDTV CODEC with ATM-Cell-Loss Compensation
IEEE Transactions on circuits and systems for video technology,
Vol. 3, No.3, June 1993.

[KOMP92] Kompella V.P. et al.
Multicasting for multimedia applications
Proc. of the Tenth annual joint conf. of the IEEE Computer and communications Societies, INFOCOM '92: One World Through communications, Florence, May 6 - 8, 1992, Technical Program Chairs: L. Fralta and J.F. Kurose, p. 2078-2085, Vol. 3

[KOU81] Kou L. et al.
A fast algorithm for Steiner trees
Acta Informatica, Vol. 15, p. 141-145
Springer-Verlag 1981

[KOU90] Kou L.
On efficient implementation of an approximation algorithm for the Steiner tree problem
Acta Informatica, Vol. 27, p. 369-380
Springer-Verlag 1990

[LEU91] Leung Y.W. and T.S. Yum
Efficient algorithms for multiple destinations routing
Proceedings of the IEEE International Conference on Communications: Communications: Rising to the heights, Denver, 1991, June 23-26, Technical Program Chairman: D. Brady, p 1311-1317, Vol. 2

[LEU93] Leung Y.W. and T.S. Yum
Efficient algorithms for multiple destinations routeing
International journal of digital and analog communication systems, Vol. 6, p. 87-95, John Wiley & Sons, 1993

[LOUI93] Loui, A.C.P. et al.
High resolution still-image transmission based on CCITT H.261 Codec
IEEE Transactions on circuits and video technology, Vol. 3, No. 2, April 1993

[MEHL88] Mehlhorn K.
A faster approximation algorithm for the Steiner problem in graphs
Information Processing Letters, No. 27, p. 125-128
Elseviers Science Publishers B.V., North Holland, 1988

[PAN94] Pancha, P. et al.
MPEG coding for variable bit rate video transmission
IEEE Communications Magazine, May 1994

[RAY86] Rayward-Smith V.J. and A. Clare
On finding Steiner vertices
Networks, Vol. 16, p. 283-294
John Wiley & Sons, 1986

[REG90] Regnier, J. and Cameron, W.H.
State-dependent dynamic traffic management for telephone networks
IEEE Communications Magazine, October 1990, p. 42-53

[SIB94] Sibal, S. and Desimone, A.
Controlling alternate routing in general-mesh flow networks
Proc. of ACM SIGCOMM, August 1994, p. 168-177

[SIM901] SIM 90/86
A comparison of coding performance between 1-layer and 2-layer codecs (BTRL)
Related COST 211 bis/ter paper

[SIM902] SIM 90/86
A comparison of variable versus constant bit rate and the impact on the picture quality for fixed frame rates (PTT/RNL)
Related COST 211 bis/ter paper

[SIM903] SIM 90/133
Comparison of VBR with CBR coding, evaluation of the statistical multiplexing gain (SIEMENS AG)
Related COST 211 bis/ter paper

[STIE91] Stieglitz, A.R. von
Design of topologies for ATM network architectures

Eindrapport RNL 1554/91, PTT Research 1991

[TAKA94] Taka M. and Abe T.
Network reliability design techniques to improve customer satisfaction
IEEE Communications Magazine, October 1994

[TAN93] Tanaka Y. and P.C. Huang
Multiple destination routing algorithms
IEICE Transactions on Communications, Vol. E76-B, No. 5 May 1993

[TINA93] An overview of the TINA Consortium Work
Document no. TB_G1.HR.OO1._1.0_93, December 1993

[TOD92] Tode H. et al.
Multicast routing algorithm for nodal load balancing
Proc. of the Tenth annual joint conf. of the IEEE Computer and communications Societies, INFOCOM '92: One World Through communications, Florence, May 6 - 8, 1992, Technical Program Chairs: L. Fralta and J.F. Kurose, p. 2086-2095, Vol. 3

[VERM93] Verma D.C.
Routing reserved bandwidth multi-point connections
Proc. of IEEE/ACM SIGCOMM 1993, p.96-105

[VRIES94] Vries R.J.F. de, Blokland A.A.E. van
ATM: A Status Report - Issue 5 -
PTT Research, R&D-RA-94-794

[WAX88] Waxman B.M.
Routing of multipoint connections
IEEE Journal on Sel. Areas in communications, Vol. 6, No. 9, December 1988

[WIN87] Winter P.
Steiner problem in networks: a survey
Networks, Vol. 17, p. 129-167
John Wiley & Sons, 1987

[WONG84] Wong R.T.
A dual ascent approach for Steiner tree problems on a directed graph
Mathematical Programming, Vol. 28, p. 271-287, 1984

[WU86] Wu J.F. et al.
A faster approximation algorithm for the Steiner problem in graphs
Acta Informatica, Vol. 23, p. 223-229
Springer-Verlag 1986

[YANG94] Yang O.W.W.
Two-centre tree topologies for metropolitan area networks
IEE Proceedings on Communications, Vol. 141, No. 4, August 1994

[ZELI93]	Zelikovsky A.Z. A faster approximation algorithm for the Steiner tree problem in graphs Information Processing Letters, No. 46, p. 79-83 Elseviers Science Publishers B.V., North Holland, 1993
[ZHA93]	Zhang, L. et al. A new resource reservation protocol IEEE/ACM Transactions on Networking 1 Semptember 1993

Appendix B Reverse point-multipoint OAM problems

As described in section 3.3.2, Operation and Maintenance (OAM) functions comprise both fault management and performance management functions for Virtual Path connections (VPCs) and Virtual Channel connections (VCCs). Fault management functions are needed to detect and signal malfunctioning of (parts of) connections. Several signals are required to inform other entities in the network of VPC or VCC failures [I610]:

- the VP/VC Alarm Indication Signal (AIS)
- the VP/VC Remote Defect Indication (RDI)
- the VPC/VCC Continuity Check (CC)

The connecting point (switch or cross-connect) detecting a fault sends downstream AIS cells. Connection endpoints receiving AIS cells sends FERF cells back in the upstream direction. Continuity Check cells are generated by the transmitting (source) endpoint when it does not send user cells. In this way, involved network entities are able to check whether a VPC or VCC is still alive.

Performance management functions are required to detect degraded performance of VPCs, VCCs or VPC/VCC segments in terms of cell losses, mutilation of cells, excessive delay of cells, etc. Monitoring results, e.g. degraded performance, are reported to the far end using the reverse OAM flow or via some other means, such as the Telecommunication Management Network (TMN). These monitoring results are generated by either connection endpoints or intermediate points. Forward performance monitoring and backward reporting is performed using performance management OAM cells (except reporting to the TMN).
Performance monitoring can be activated either during connection establishment or at any time after the connection has been established. Such activation (and associated deactivation) is initiated by the TMN.

Both the fault management and performance management OAM functions can either be used by end-to-end or segment OAM flows. If segment OAM flow is performed, the segment in consideration is responsible for removing its own OAM cells at the segment endpoints.
OAM flows can be activated in forward and backward directions independently. For example, the Continuity Check can be end-to-end in forward direction and based on segments in backward direction. Virtual Channel and Virtual Path connections can be partitioned in segments independently.

Additional ATM Layer OAM functions are currently under study. The OAM cell *loopback capability* is one such function. The ATM Layer loopback capability allows for operations information to be inserted at one location along a virtual (path/channel) connection and returned (or looped back) at a different location, without having to take the virtual connection out-of-service. This capability is performed by non-intrusively inserting a loopback OAM cell at an accessible point along the virtual connection (i.e. at a local end-point or intermediate point) with instructions in the OAM cell payload for the cell to be looped back at one or two other identifiable points.
Potential uses of the loopback capability are:

- Pre-service Connectivity Verification
- VPC/VCC Fault Sectionalization
- On-demand Delay Measurements

Reverse OAM problems

Bidirectional employment of OAM flows in point-multipoint connections introduces two problems for the reverse direction:

- For each received OAM cell of the reverse OAM flows, the root is not able to identify the sending leaf.
- The bandwidth of the reverse OAM flows at the root is the sum of the bandwidth of the reverse OAM flows at all the leaves. In case of many leaves the reverse bandwidth required at the root can be very high. The same happens for processing power required at the root for processing the reverse flows.

These two problems are currently subject of international standardization efforts (ETSI NA 5, ITU-T Study group 13). A complete solution for these problems is not available yet. Nevertheless, under some conditions, reverse OAM flows cause no problems if one accepts only a less precise assessment of the performance of the connection in consideration. These conditions are [CSELT95]:

- Each leaf is considered to be essential to the connection; a fault on the connection, no matter how many leaves are affected, makes the complete p-mp connection unavailable. The root knows the number of connected leaves.
- The number of leaves is such that the root can receive and process all reverse directed OAM cells. Reduction in precision of assessment of connection performance is acceptable.

Point-multipoint connections under these conditions are also called *conference like* connections. The reduced number of leaves enables the use of *leaf identifiers*, although still a point of discussion within the standardization organisations, it is already proposed for signalling.

The conclusion is that the problems remain for connections with a large number of leaves, like those supporting the class of distributive services. Because reverse OAM problems are in fact scaling problems (connections with a reduced number of leaves do not suffer), a practical solution should be based on partitioning (segmentation) of the p-mp connection. As described, OAM segments are supported by Rec. I.610. Figure I illustrates how segmentation can be applied to p-mp connections. Figure I shows the copying and merging functions inside a Network Element (e.g. switch or cross-connect), also the possible OAM sink locations and OAM source locations are illustrated. The connecting points illustrate that OAM segments cannot overlap.

The solution given here presumes that the nework is managed by a TMN; Network Elements (NE) have either a TMN-interface or communicate through TMN-interfaces of neighboring NE's.

Continuity Check

In p-mp direction, end-to-end Continuity Check functions can be used, so segment Continuity Checks are not required.
In mp-p direction, end-to-end CC function causes concentration problems at the root sink point. Therefore, instead segment CC function should be enabled in this direction. Segment endpoints generate CC cells, the root (OAM sink) of the segment terminates CC cell streams.

AIS

The connecting point detecting a fault sends downstream end-to-end AIS cells and possibly segment AIS cells. Also the TMN is reported by an alarm (which includes a NE identifier).

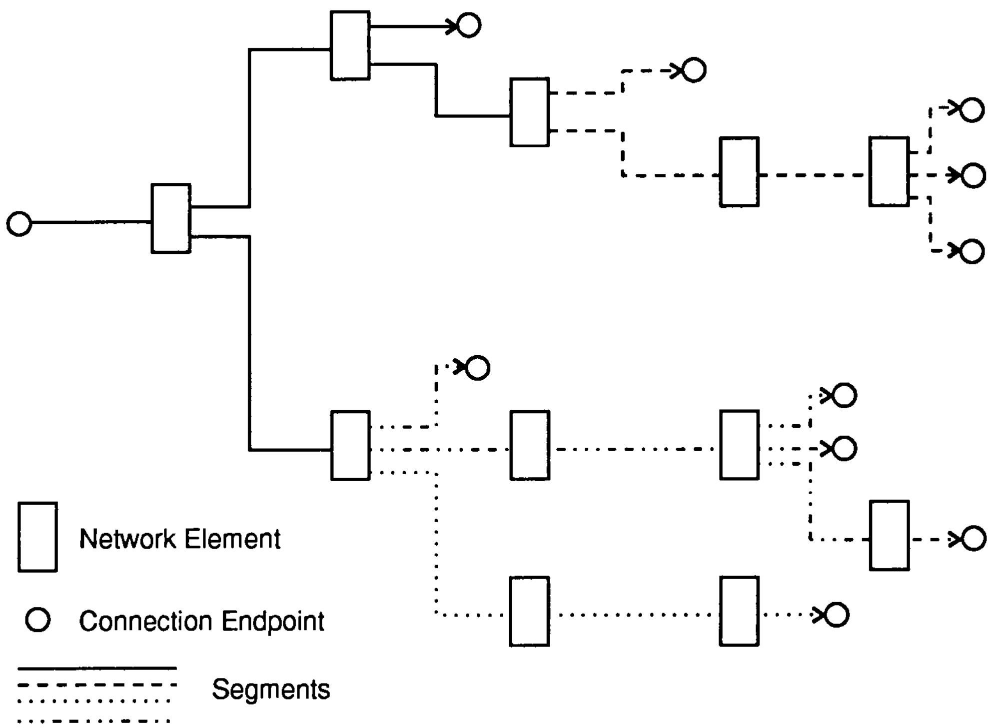


Figure I OAM segments in a p-mp connection

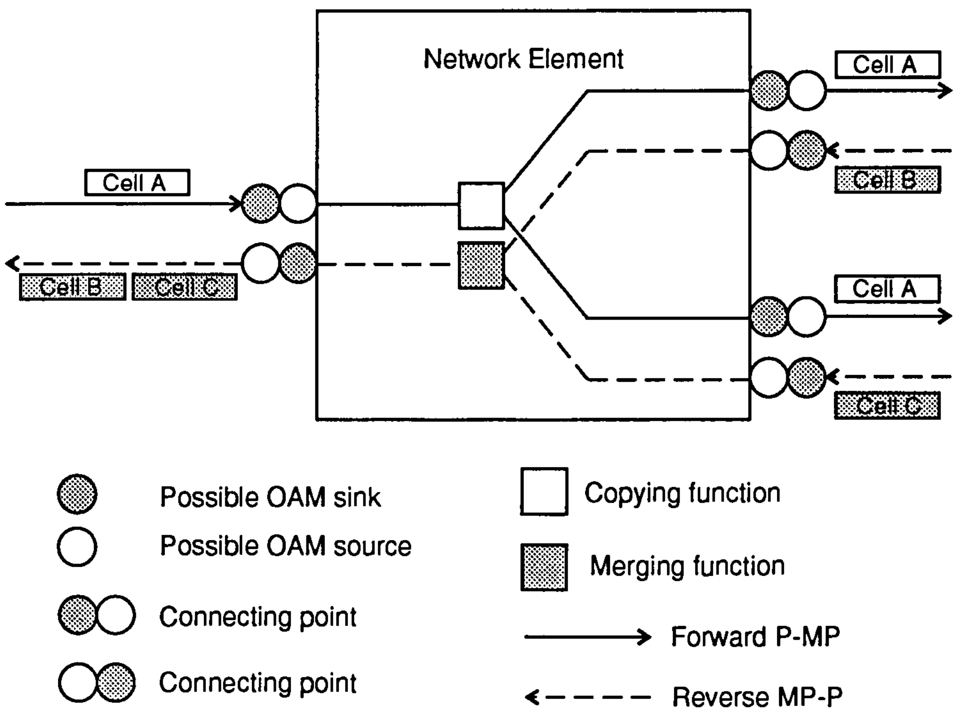


Figure I Branching point

Connection endpoints react by sending end-to-end RDI cells. These end-to-end RDI cells cause concentration problems. A possible solution is to have the possibility of disabling these backward directed OAM flows and enabling backward segment RDI cells. This solution requires an adaption of recommendation I.610. Connecting points receiving RDI signals report this to the

TMN in case they have a TMN interface. The TMN can respond by disabling these alarms in case of a large number involved connection points per Network Element (e.g. a LEX in case of a broadcast connection)

In this way, RDI signals are terminated at the OAM sinc of the segment where the failure was detected. If further upstream transport of the RDI signal is desired, some communication mechanism is required between neighboring OAM segments.
Based on the alarms and the locations where they are generated, the TMN is able to (globally) localize the fault. Further precision fault sectionalization can be achieved with the loopback function.

Loopback

Loopback cells are either returned at one or two specified points along the virtual connection (identified by the *Loopback Location ID Fields*) or at the receiving connection end-point (the *Loopback Indication Field* identifies whether the OAM cell is to be looped back at the end-point).

Loopback cells sent by a connecting point reach all connection end-points downstream, except those in the subtree rooted by the return point. Therefore loopback cells should not be returned at multipoint connection end-points (indicated by the Loopback Indication Field) unless the endpoint is included in the Loopback Location ID Fields.

When loopback cells are used for precise VPC/VCC fault sectionalization in p-mp connections, the TMN initiating these loopback cells was already able to localize the OAM segment of the fault based on the alarms sent by NE's involved, as described above. More precise fault sectionalization can now be achieved by initiating loopback cells to be returned at connection points in that OAM segment. (Note: only one loopback cell is expected in upstream direction for each loopback cell sent in downstream direction, although the downstream directed cell will be copied and transported through the whole underlying subtree.)

Performance monitoring and reporting

Performance monitoring OAM cells can be enabled in forward direction using end-to-end flow. Intermediate nodes are able to generate downstream directed AIS cells when the performance falls below certain tresholds or report to the TMN. Receiving end-points are able to generate segment RDI cells when the performance falls below certain tresholds or report to the TMN.

Appendix C Additional ATM issues

The B-ISDN Protocol Reference Model

The B-ISDN Protocol Reference Model (B-ISDN PRM), a protocol framwork for the standardisation of B-ISDN/ATM that has been standardized within ITU-T, is shown in figure 3.4 [I311].

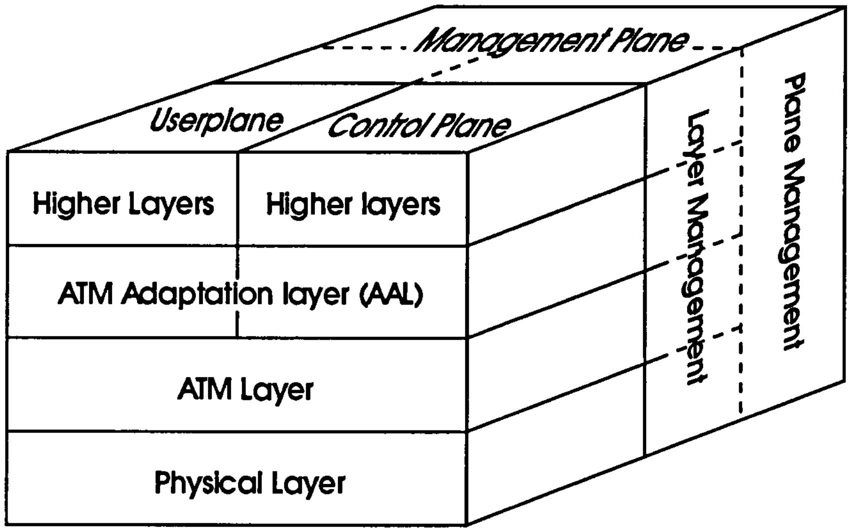


Figure I B-ISDN Protocol Reference Model

The B-ISDN PRM is in essence composed of th Physical Layer (PL), the ATM layer, the ATM Adaptation Layer (AAL), and three planes, the user plane , the control plane and the management plane. Above the Physical Layer, which provides the transfer of a stream of ATM layer cells over a medium, the ATM layer provides service independent cell transfer through the network, and the AAL provides higher level service dependent functions. The upper layers in the Control and Mangament planes provide Call Control, Connection Control and Network Supervision functions.

Physical layer

The main function of the Physical layer is to provide the conveyance of ATM layer cells, either assigned or unassigned cells, between the ATM layers of geographically separated network elements such as network nodes and terminal equipment. It consists of two sublayers performing different functions.

Physical Medium Sublayer

- medium dependent bit functions
- correct transmission and reception of bits, including bit timing

Transmission Convergence Layer

- Cell rate decoupling
- HEC header sequence generation/verification
- Cell delineation
- Transmission frame adaptation
- Transmission frame generation/recovery

ITU-T recomments a 155.52 Mbit/s bi-directional interface between user and network. Also a 622.08 Mbit/s interface is recommended, either symmetrical or asymmetrical with a rate of 155.52 Mbit/s in the upstream direction.

ATM Layer

The ATM Layer is common to all ATM layer service users including signalling and systems management. The basic functions that are associated with the ATM layer are:

- Cell switching and VCI and/or VPI translation
- Cell multiplexing and demultiplexing
- Cell header generation and extraction
- Generic Flow Control (only at the UNI)

ATM Adaptation Layer

The service provided by the ATM layer, i.e. the transfer of information fields from a source to one (or more) destinations, is a generic (service independent) transfer service which cannot directly be used by current higher layer protocols. Therefore, an intermediate layer between the ATM layer and the higher layers is needed: the ATM Adaptation Layer (AAL). The main purpose of this layer is to isolate the higher layers from the specific characteristics of the ATM layer by mapping in some manner the higher layer PDUs into the information fields of the ATM cells and vice versa.
The AAL consists of two sublayers: the Segmentation And Reassembly (SAR) sublayer for segmentation and reassembly of the higher layer PDUs and the Convergence Sublayer (CS). The CS is positioned directly below the higher layers and thus is responsible for the AAl service offering at the AAL Service Access Points (AAL-SAPs). The CS is inherently service dependent. To prevent uncontrolled growth of service offering and to minimise the number of protocols used in B-ISDN four *service classes* have been defined and standardized. The three key parameters, that have been used to derive the service classes are:

- **Timing Relation**
  A timing relation between source and destination(s) can be required or not required.
- **Bit Rate**
  The bit rate can be constant or variable
- **Connection Mode**
  Two connection modes can identified: connection-oriented and connectionless.

Table A shows the definition of the service classes, i.e. the relation between the classes and the key parameters.

Table A. Service classification for the AAL

	Class A	Class B	Class C	Class D
Timing relation between source and destination(s)	Required		Not required	
Bit rate	Constant	Variable		
Connection mode	Connection-oriented			connection-less

AAL types

ITU-T/Recommendation I.361 describes a number of AAL types (protocols) which consist of combinations of SAR and CS functions and can support higher layer services belonging to one of the defined classes (A through D). Up to now four AAL types have been elaborated in ITU-T: AAL type 1, AAL type 2, AAL type 3/4 and AAL type 5. AAL type 1 is directed to class A services, AAL type 2 to class B services and AAL type 3/4 to class C and D services. AAL type 5 has been included recently for high speed data services.

Broadband interfaces

The B-ISDN Reference Configuration

Figure IV shows the B-ISDN reference configuration with its functional groupings (B-TE1, B-TE2, B-NT2, and B-TA) and reference points (R, S_B and T_B).

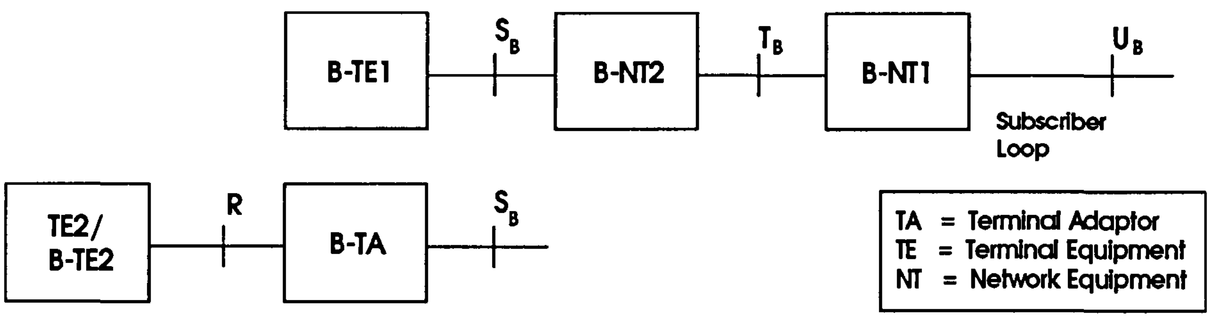


Figure I General B-ISDN reference configuration

The Broadband Network Termination 1 (B-NT1) is positioned between the T_B reference point and the U_B reference point at the local exchange. A second Broadband Network Termination (B-NT2) is located between the S_B and T_B reference points.

The User-Network Interface

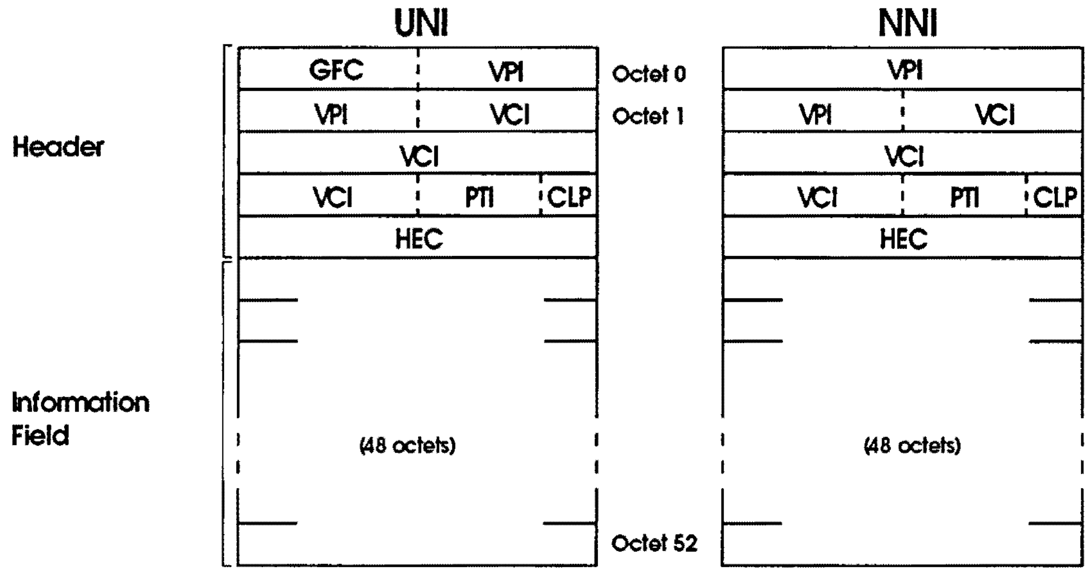
According to the ITU-T definition of the UNI is the interface between terminal equipment and a network termination at which the network access protocols apply. The structure of the cell used at the UNI differ slightly from the cell structure at NNI, see figure 3.5.

The Network Node Interface

The NNI is the standardized interface between network nodes. The NNI has not been specified by ITU-T yet but maximum commolality between functions of the Physical Layer at the UNI and corresponding functions at the NNI is aimed at.

Cell structure

The cell structures of ATM cells at UNI and at NNI differ. They are shown in figure V.



GFC	: Generic Flow Control	4-0 bits	PTI	: Payload Type Identification	3 bits
VPI	: Virtual Path Identifier	8-12 bits	CLP	: Cell Loss Priority	1 bit
VCI	: Virtual Channel Identifier	16 bits	HEC	: Header Error Control	8 bits

Figure I ATM cell structure at UNI and NNI

User plane

The user plane provides the user information flow transfer, with associated controls (e.g. flow/error control), verifications and re-transmissions if necessary. The network elements involved in the transport of user plane information are:

VP cross-connect:	connects VP links and translates VPI
VC cross-connect:	connects VC links, terminates VPCs and translates VCI values
VP-VC cross-connect:	acts both as VP and VC cross-connect
VP switch:	connects VP links and translates VPI values
VC switch:	connect VC links, terminates VPCs ad translates VCI values
VP-VC switch:	acts both as VP and VC switch.

The cross-connects are directed by management plane functions and switches are directed by control plane functions.

Control plane

This plane handles the call control and connection control information. It deals with the signalling flow necessary to set-up calls and connections, to re-negotiate call or connection characteristics, to terminate calls and to release connections.

Management plane

Two groups of functions can be distuingished within the management plane, namely Layer Management functions and Plane Management functions.

Plane Management functions will satisfy the need for communication between the user, control and management planes - it provides co-ordination between all the planes. Plane management relates to a system as a whole and has no layered structure.

Layer Management performs functions related to resources and parameters residing in protocol entities of the system. Layer mamagement has a layered structure. For each layer it handles the related Operation and Management (OAM) information flows.

Appendix D Video compression techniques

The efficient real-time transport of broadcast-quality video images requires compression techniques to reduce the bit rates and lower the sensitiveness for information losses. Several coding algorithms are proposed for constant and variable bit rate video. This variety in coding techniques has made the task of characterizing video sources difficult. However, some video coding algorithms have already been standardized or are currently being standardized. These will be introduced here.

H.261

The H.261 coding algorithm was developed for use with STM networks [H261]. Thus, it was optimized to operate at fixed bit rates, irrespective of picture content. The buffer regulation process is such that only a small variation of the instantaneous bit rate is permitted. H.261 used with STM networks achieves variable picture quality and constant bit rate. The video coding (and decoding) methods use bit rates of p * 64 kbit/s, where p is in the range of 1 to 30. H.261 covers the use of 625- and 525-line television standards by defining a common intermediate format (CIF). The source coder operates on pictures based on this CIF.

MPEG

An experts group in ISO, known as the Moving Picture Experts Group (MPEG), is developing coding standards for generic applications, initially in particular for storage applications. Currently this group is working on two standards known as MPEG1 and MPEG2, where the former is already used in some applications (e.g. CD-Interactive). The principles behind the MPEG coding techniques are described below [COST93].

MPEG1

This is a standard for coding progressive images, primarily those of resolution and frequency similar to the Common Intermediate Format defined in H.261. It has been optimized for (constant) bit rates between 1.1 Mbit/s and 1.5 Mbit/s. It has become an international standard: ISO/IEC 11172.

The basic coding algorithm is similar to H.261. A major difference is the definition of three picture types: intra (I), predicted (P) and interpolated (B) pictures, where the intra pictures form the basic pictures. A typical group of pictures is depicted in figure I.

Interpolated pictures occur in a different order in the coded bit stream to the display order. This results in a reorder delay which depends on the number of interpolated pictures between intra and predicted pictures. The number of interpolated pictures between intra and predicted pictures can be zero or any positive value. The use of interpolated pictures may incur too much delay for face to face applications; such applications may just use intra and predicted pictures. Because the MPEG1 coding is optimized for constant bit rates between 1.1 Mbit/s and 1.5 Mbit/s it will not efficiently support higher bit rate video encoding which is required for extended quality video services.

Thus, if MPEG1 coding is utilized for video information transmission, one constant bit rate channel is required at 1.5 Mbit/s.

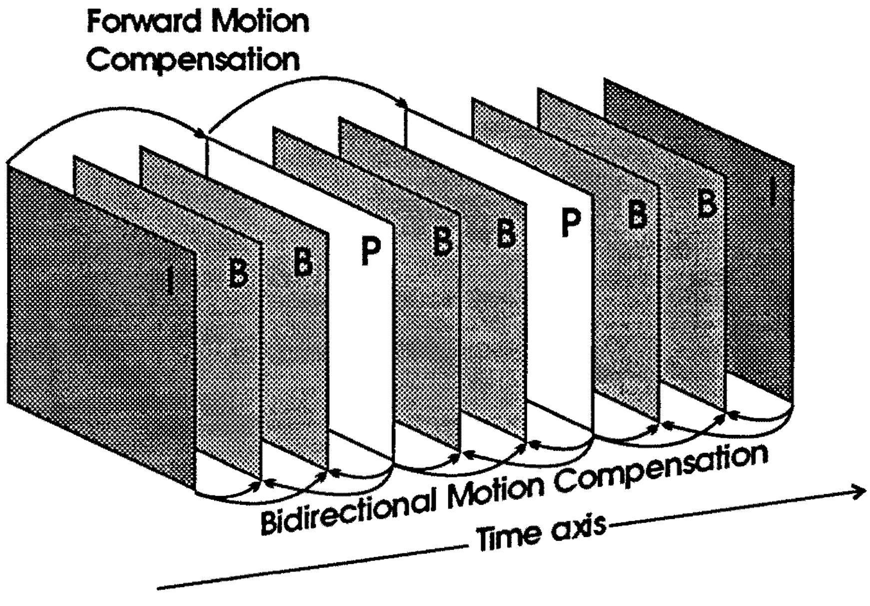


Figure I MPEG Group of Pictures structure

MPEG2

This is a standard for coding interlaced (but also progressive) images, primarily those of resolution and frequency similar to that defined in CCIR recommendation 601 (720x480/30/2:1 or 720x576/25/2:1). However, this standard will also support extended picture quality formats such as Extended Definition Television (EDTV) or High Definition Television (HDTV). The standard is being optimized in the MPEG group for bit rates between 4 Mbit/s and 9 Mbit/s. The basic coding algorithm will be similar to MPEG1, but is still under development. It is likely that the standard will be used for applications ranging from digital Video Tape Recording (VTR) to broadcast TV and even broadcast HDTV, which requires 10 - 20 Mbit/s [COST93].
MPEG-2 has attracted unprecedented industrial interest, and will probably play a dominant role in the future of digital television. MPEG-2 will be a "tool-kit" with a special syntax to specify the various bit streams for video representation. As a generic International Standard, MPEG-2 Video is being defined in terms of extensible profiles, each of which will support the features needed by an important class of applications. The highest level of picture quality defined by MPEG-2 corresponds to HDTV resolutions. The definition of this hierarchical structure provides upward and downward compatibility; forward and backward compatibility (with MPEG-1) is also a main objective.

The layered approach enables the standard to meet the requirements of video service interworking. Several proposals are made for implementation of the layered coding mechanism; figure I shows a generic scheme for these proposed encoders.
In the layered approach the first, or basic, profile provides the basic pictures (low quality, low codec complexity). Each higher level profile could add some (standardized) extra picture components which increases the picture quality. The information of these distinct layers can be transported in several manners. If a two layered codec is applied for example, then both bit streams can be separated using two priority levels, while using only one channel. On the other hand, if multiple layers are used, several channels could be used to transport the multi-layered bit stream.

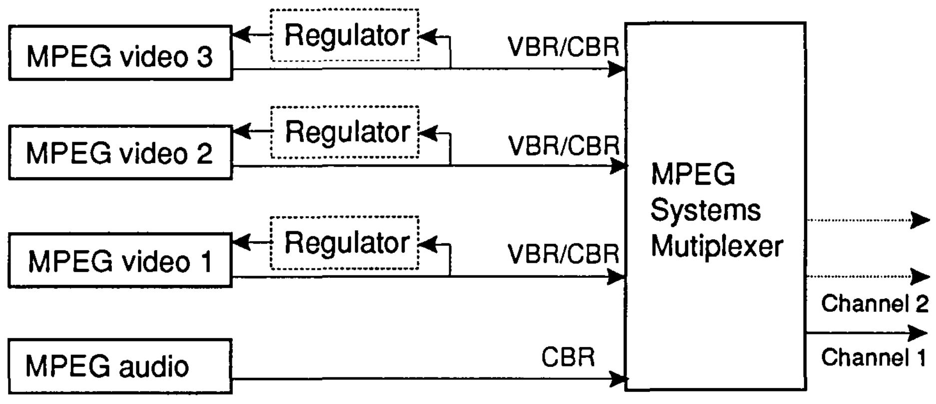


Figure 1 Generic scheme of MPEG2 Systems

MPEG is also developing the MPEG-2 Audio Standard for low bit rate coding of multi-channel audio, as well as the MPEG-2 Systems Standard, which is of particular importance, because it specifies coding formats for multiplexing audio, video, and other data into a form suitable for transmission or storage. With the addition of the MPEG-2 Systems component, MPEG-2 becomes a powerful and flexible multi-purpose coding method which enables a range of possibilities for the transport of the video signal. These include the multiplexing of multi-layered coding bit streams into one channel, or the provision of a separated channel for HDTV quality for example.

The flexibility of MPEG-2 coding systems makes them appropriate canditates for use in telecommunication networks. However, the same flexibility regarding the required number of channels, the transport mode (VBR or CBR) and the required bandwidth makes a proper estimate about these issues a difficult matter. It is more likely that future MPEG-2 coding systems will depend on specific network constraints.

Effect of coding and decoding delay

The coding and decoding delay times in MPEG Video coding depend on the coding mode in use. There are three coding modes: intra (I), predictive (P) and interpolated (B). The (bidirectional) interpolated mode introduces the longest delay because the prediction is based on pictures from past and future frames. Assume that M - 1 B frames are inserted between two adjacent P frames, then the total encoding delay, including preprocessing, is M + 1.3 frames [CHEN93, KAW93]. A typical value for M is 3, which would imply an encoding delay of approximately 110 ms.

The decoding delay is about a half-frame delay, necessary to convert the signal into an interlaced one (one frame contains two fields in the interlaced mode).

This analysis makes clear that the encoding and decoding delays are not relevant to distributive services, in contrast to interactive services where the requirements on delay times are much more critical (for instance: < 200ms).

Note

Encoding and decoding delays for video in distributive services can be of significant importance in the case some interactive features are added to the service. For example in TV programs where the user can interact with certain games or other distributive entertainment.

Appendix E Example of a flooding algorithm

A distributed control, strict local information, static flooding algorithm is described here. The algorithm is described in detail in [JIA93]. Packet exchange is the basic communication method employed in the distributed algorithm.
The communication takes place mainly between immediately neighboring nodes. There are seven types of control packets used in the algorithm, which are:

- A *TRY* packet contains the current source node number, a list of destination nodes, (the minimum bandwidth required for the tree?), parameters of an available path from the source to this node, P_{try} (including the cumulative cost), and the maximum transmission range (i.e. the age) of the packet. It is the major means used by a node trying to make the multicast tree grow to reach all of the destination nodes.
- A *CONNECT* packet contains the source node number, P_{try} and the source node number from which the packet originates. It is the response sent by destination node sent to the source node.
- A *DISCONNECT* packet only contains the source node number. It is used by a non-destination node to remove itself from the tree (as an unnecessary leaf) and to inform the corresponding adjacent nodes.
- A *CHECK* packet contains the source node number and the list of destination nodes. It is sent out by the source node to stabilize the tree when it is found and prvent further changes to the path.
- A *CONFIRM* packet contains the same types of information as those in a *CONNECT* packet. It is a response from a destination node to the source node to confirm the path and the available bandwidth.
- A *DISMISSED* packet contains the source node number only. It is broadcast from the source node to announce the failure of a path finding process.
- A *COMPLETED* packet contains the source node number and the available capacity of the tree. The source node announces the succesful completion of a tree search by broadcasting this packet.

The *TRY* packets may cause a broadcast storm due to their widespread transmission, and thus need an age field to prevent it. The other types of packets are either unicast to single destination (the source node or a neighbouring node) or multicast through a fixed route.
During a tree searching process, each node must maintain the following information:

- the incoming node, which is the parent node of this node in the tree
- an outgoing table, which lists all the links connecting child nodes on the tree to this node
- a set of parameters, $P_{current}$, describing the available path from the source node to this node.

The algorithm is initialized by a source node starting the *source node procedure*. It involves three stages: *tree building, confirming and concluding*. A tree building stage begins with the source node broadcasting a *TRY* packet. The broadcast triggers other nodes to find a capable path back to the source node. When all of the destination nodes report the finding of such a path, by sending *CONNECT* packets to the source node, the tree building stage end succesfully. However, it is possible that a capable Steiner tree does not exist. In this case, none of the nodes know the situation because after the propagation of *TRY* packets they either wait for replies or are not contacted at all. This can be detected by a timer at the source node.

The confirming stage begins with the source node sending out *CHECK* packets, and ends by all of the destination nodes replying with *CONFIRM* packets. While a *CONFIRM* packet travels back to the source, it stabilizes path.
The concluding stage is to inform the network of the result of the path finding process by broadcasting the *COMPLETED* packet. In the event of a time-out from the timer, there is no confirming stage, and a failure report (i.e. a *DISMISSED* packet) is sent to conclude the algorithm.

The computational complexity of the distributed path finding algorithm can be analyzed stage by stage. Since the processing at each node involves simple calculations and selections based on its local information, the computational complexity is proportional to the maximum length of paths along which packets travel.
At tree building stage, the longest path from which a destination node may be reached by a *TRY* packet consists of all of the links in the network. Since the maximum number of nodes is $1/2(|N|^2 - |N|)$, this is also the length of such a path. If the path happens to be a viable one, a *CONNECT* packet is generated, and travels along the same route back to the source. Therefore, the worst case performance is $O(|N|^2)$. The complexity of the confirming stage and the concluding stage is the same, due to the identical request-response cycles.

Note
The computational complexity applies here to the number of messages between adjacent nodes.