MASTER

The use of ATM in an inhouse radio network

Maessen, R.B.

Award date: 1992

Disclaimer
This document contains a student thesis (bachelor's or master's), as authored by a student at Eindhoven University of Technology. Student theses are made available in the TU/e repository upon obtaining the required degree. The grade received is not published on the document as presented in the repository. The required complexity or quality of research of student theses may vary by program, and the required minimum study period may vary in duration.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
THE USE OF ATM
IN AN INHOUSE RADIO NETWORK

by

R.B. Maessen

Eindhoven, 19th june, 1992

Report of graduation study
performed from September 1991 until June 1992
Supervisor: prof. J. de Stigter
Coach: ir. P.F.M. Smulders

The faculty of Electrical Engineering of the Eindhoven University of Technology does not accept any responsibility regarding the contents of graduation reports.
Abstract

Since 1988 the research programme of the Telecommunications Division of the Eindhoven University of Technology has contained a project in which various aspects of indoor radio have been studied. The objective of this project is to examine the feasibility of an indoor wireless broadband network and to produce a design for such a system.

The use of radio links in local area networks increases. This has a couple of reasons. First, a radio network is more flexible than a wired network. Terminals can be moved through a building without difficulty. This has the advantage that network services are available for each user where and when he wants them. Second, a radio network can be installed easily since there is no necessity of difficult and expensive wiring.

Until now most indoor radio networks are only suitable for smallband services, like speech and low speed data. In the future, users will need networks that can transport higher bitrate services. This makes a broadband radio network desirable.

The estimation of the chances for realising an Indoor Radio Integrated Broadband Network (IRIBN) in the 60 GHz band using the Asynchronous Transfer Mode (ATM) has been the subject of this feasibility study. An IRIBN based on ATM has been described. The multi access protocols that can be used in an IRIBN have been reviewed. The Improved Idle Casting Multiple Access Collision Detect Protocol (I-ICMA-CD) is chosen and analyzed. A protocol for the information transfer and channel allocation has been suggested.

Emphasis should lay on the fact that the radio links are still studied. Their qualities are still uncertain. Especially in the area of the possible bit rates, the packet detection delay and the Bit Error Rate.

The realisation of the IRIBN looks to be possible. A (technical) proposal has been done for such a network. The economical aspects have only been briefly discussed. The use of ATM implies that the network will be compatible with the future world standard for broadband networks. The costs of the system have not yet been calculated.

During the study it became clear that the currently developed indoor radio networks, are still pre ISDN networks. The exchange of the IBN for an ISDN can be an interesting option. The proposed radio system can still be used then, but the requirements will be less.

A lot of further study remains to be done. Inter room networks have not been studied. This is of particular interest for commercial networks. Further, the queuing model of the network can be searched for.
Contents

1. Introduction ......................................................... 1

2. IBN in an indoor environment ................................. 2
   2.1. Introduction in integrated broadband communications networks 2
   2.2. Asynchronous Transfer Mode ................................ 2
   2.3. IBN in an indoor environment .............................. 3
   2.4. Indoor radio ATM ............................................. 4
   2.5. Indoor ATM-based IBN protocol reference model ............ 5
   2.6. Conclusions ................................................ 6

3. The indoor ATM-based IBN ........................................ 7
   3.1. General description of the network ....................... 7
   3.2. Network configuration .................................... 7
   3.3. Access to the network .................................. 10
   3.4. Channel allocation ....................................... 11
   3.5. Conclusions ................................................ 13

4. The multi access protocol ...................................... 14
   4.1. Tasks of the protocol ................................... 14
   4.2. The protocols which can be used ......................... 14
   4.3. SRMA ......................................................... 15
      4.3.1. Operation of SRMA .................................. 15
      4.3.2. Analysis of SRMA .................................... 16
      4.3.2.1. Delay $D_1$ ....................................... 17
      4.3.2.2. Delay $D_2$ ....................................... 18
4.4. Request channel multi-access protocols .................................. 19
4.4.1. Busy Channel Multi Access (BCMA) ................................. 20
4.4.1.1. Performance of BCMA ........................................... 21
4.4.2. Busy Tone Multi Access (BTMA) .................................... 22
4.4.3. Idle Casting Multi Access (ICMA) .................................. 23
4.4.3.1. ICMA-CD ............................................................. 24
4.4.3.2. Improved ICMA-CD (I-ICMA-CD) .............................. 26
4.4.3.3. Recent ICMA-variants ............................................ 27
4.4.4. Polling ................................................................. 27
4.4.5. Relations between SRMA and Polling .............................. 29
4.5. Conclusions ........................................................................ 30

5. The Improved ICMA Collision Detect protocol ...................... 31
5.1. Protocol description and functions ..................................... 31
5.2. The protocol scheme ....................................................... 32
5.3. Protocol analysis ............................................................ 34
5.4. Protocol performance with transmission errors .................. 36
5.5. Conclusions .................................................................... 39

6. Protocol for information transfer and channel allocation ........ 40
6.1. The existing channel allocation protocol ......................... 40
6.1.1. Protocol description ................................................... 40
6.1.2. Cell delay ............................................................... 41
6.1.3. Evaluation ............................................................... 41
6.2. Network properties ....................................................... 42
6.2.1. Network performance ................................................ 42
6.2.2. Statistical and deterministic multiplexing ...................... 42
6.2.3. Traffic characteristics ............................................... 43
6.3. The basic allocation protocol .......................................... 45
6.4. The information channels ............................................... 46
6.5. The error control .......................................................... 48
6.5.1. The local error control of the ATM cells ....................... 48
6.5.2. The error control of the control tokens ......................... 49
6.5.3. Choice of total error control ...................................... 50
6.6. The token system ........................................ 50
  6.6.1. Token System on the links ........................ 51
  6.6.2. Class A traffic .................................... 52
  6.6.3. Class B traffic .................................... 53
  6.6.4. Class C & D traffic ................................ 54
6.7. Control of token mechanism .............................. 55
  6.7.1. Determination and description of the system state 55
  6.7.2. Capacity allocation .................................. 57
  6.7.3. Signalling, which priority? .......................... 59
6.8. Queueing analysis ........................................ 60
  6.8.1. M/D/1 queueing system ............................. 60
  6.8.2. TDMA queueing model ............................... 65
6.9. Conclusions .............................................. 67

7. System Proposal ........................................... 68

8. Conclusions ................................................. 69

References ..................................................... 70

List of abbreviations ......................................... 75

List of variables ............................................. 77
1. Introduction.

Since 1988 the research programme of the Telecommunications Division of the Eindhoven University of Technology has contained a project in which various aspects of indoor radio have been studied. The objective of this project is to examine the feasibility of an indoor wireless broadband network and to produce a design for such a system.

The use of radio links in local area networks increases. This has a couple of reasons. First, a radio network is more flexible than a wired network. Terminals can be moved through a building without difficulty. This has the advantage that network services are available for each user where and when he wants them. Second, a radio network can be installed easily since there is no necessity of difficult and expensive wiring.

Until now most indoor radio networks are only suitable for smallband services, like speech and low speed data. In the future, users will need networks that can transport higher bitrate services. This makes a broadband radio network desirable.

Chances for realising an Integrated Broadband Network with radio links in the 60 GHz band and the Asynchronous Transfer Mode (ATM) are estimated in the feasibility study described in this report. Chapter 2 will describe what an Integrated Broadband Network (IBN) in an indoor environment should look like. Chapter 3 deals with the ATM-based indoor radio IBN. In chapter 4 the multi access protocols that can be used in the IBN will be reviewed. The Improved Idle Casting Multiple Access Collision Detect Protocol (I-ICMA-CD) will be analyzed in chapter 5. The protocol for the information transfer and channel allocation is presented in chapter 6. A system proposal is done in chapter 7.
2. IBN in an indoor environment

2.1 Introduction to integrated broadband communications networks

The Integrated Services Digital Network (ISDN) has been developed during the past decade and is currently being implemented. ISDN is designed to narrowband services, so it is often referred to as N-ISDN nowadays. Now, a broadband successor to ISDN is under study, to transport e.g. High Definition Television (HDTV) and large files efficiently.

Main objective of this Integrated Broadband Network (IBN) is to enter the next stage of the development of telecommunications networks. This is the reason why a network concept has been chosen that will support all imaginable future services. CCITT is developing such a network.

The traffic through the IBN is subdivided according to three variables: variable or constant bit rates, necessity of message (part) arrival, and connection oriented or connectionless services. With three independent variables eight combinations are theoretically possible. Until now, the CCITT has assumed that the network has to deal with four kinds of traffic. This shows the need for a network that is as flexible as possible, that can be used by as many future services as possible.

2.2. Asynchronous Transfer Mode (ATM)

The CCITT has chosen ATM as basis for the IBN. ATM includes both switching and transfer of the data. In ATM information is only transferred when a user asks for it. The opposite of ATM is STM, the Synchronous Transfer Mode. In STM there is periodical information transfer between users, independent of their needs. Both transfer modes use fixed length information words, in STM they are called packets and in ATM cells.

ATM includes a new switching technique. As far as system simplicity and bit rate flexibility are concerned, ATM switching is situated between the classical Circuit Switching (CS) and Packet Switching (PS), as can be seen in figure 2.1. ATM tries to use the advantages, i.e. simplicity and flexibility, of both switching techniques.

In ATM asynchronous means that the users bit rate is independent of the networks bit rate. This makes the network suitable for future unequal bit rate services. This in contrast to STM, in which both bit rates are coupled. This implies that the user always has to adapt his bit rate to the network. Example of such a system is Time Division Multi Access.
The CCITT ATM reference model is given in figure 2.2.

As can be seen, ATM is located in the three lowest layers of the protocol reference model. The ATM Adaptation Layer (AAL) adapts higher layer services to ATM and vice versa.

2.3. IBN in an indoor environment

An indoor IBN will differ from a public IBN. A network within a building will always cover a very limited area and thus the distances will be relatively short. The switching capacity of in an indoor network will also be relatively simple (because of the very limited size). In a way this application can be seen as a broadband Private Automatic Branch eXchange (PABX).

Such local networks can be established using radio and/or wire links. The choice depends on what a user wants. In this feasibility study chances for realising an IBN with radio links have been calculated. These links in the 30 GHz. band or higher are currently being investigated by COST 231/WG 3 in which the Telecommunications Division of the Eindhoven University
of Technology anticipates. The radio signals in these frequency bands can not penetrate walls. This means that the use of the links is always restricted to one room, which becomes thus the basic communications cell. When it is possible to develop a network that works in a room of variable size, it will be relatively easy to extend the network to a whole building connecting the room networks with wires. This is a local wired IBN, which is beyond the scope of this study.

Since all rooms are different, a room model is needed, taking into account the worst possible complications in the sense of radio communications networks, i.e. many terminals with a lot of electromagnetic obstructions separating them and a relatively high bit error rate (BER). The bandwidth of the radio channels is of major importance, and the price of the bandwidth. The limit to the number of terminals is taken to be 200 [Sonnemans, 1990]. This worst case model isn’t necessary the basis for the network specification. In many cases a less stringent specification will have good results in practice. The above mentioned channel properties in a certain room are still subject of study by P. Smulders, the present feasibility study will give a general approach.

2.4. Indoor radio ATM

As mentioned, the CCITT has chosen ATM as the transfer and switching technique to be used in the future IBN. This is of course a very strong argument for using ATM also in the indoor IBN. Especially from a commercial point of view it is absolutely necessary for it to be compatible with this future world standard. But is this possible?

ATM has been developed for high quality communication links, like optical fibers. This made it possible to use error control on an end to end basis only, instead of the link by link error control of the classical networks. Using radio links requires some additional error control.

Advantage of packet switching over circuit switching is that in packet switching all resources are shared by the users as much as possible. This principle is also used in ATM. Using radio links, bandwidth will be scarce and thus expensive. Scarce doesn’t mean that there is not enough free bandwidth available in the 60 Ghz. band, but that using extra bandwidth requires a lot of extra hardware. Sharing will thus be necessary, and ATM incorporates this feature. In ATM it has not yet been determined what a Customers Premises Network (CPN) (ATM slang for a PABX) will look like. In the ATM users network-node cell definition four bits have been reserved for generic flow control. This mechanism takes care proper functioning of the CPN, and avoids network congestion in case of heavy network load. It has to operate
independent of the physical layer [van der Linden, 1992]. This indicates that the GFC protocols as proposed, probably will not be suitable to overcome adaptation problems with a non-standard ATM transfer medium like a radio channel. Thus it is not easy to answer the question if ATM can be used in local radio IBN. However, since ATM seems to become the IBN standard the indoor radio IBN should be based on ATM.

2.5. Indoor ATM-based IBN protocol reference model

During the development of ATM compatibility problems arose between the existing systems and the proposed ATM system. This was also the case with the transfer techniques applied to the physical links. It was decided that for the transfer of data (Physical Layer) any transfer mode can be used. Example is the fully STM oriented Standard Digital Hierarchy (SDH), which is becoming widely accepted as a broadband carrier. It has to be emphasized that all the (optional) different Physical Layers have to offer the same services to the ATM layer. With this in mind, the differences between radio based ATM network and standard ATM should best be concentrated in the Physical Layer. This is obvious because a different physical link is used. At the ATM layer level the indoor network must be exactly the same as in standard ATM. In other words, the user of the indoor network can communicate normally over the ATM layer with other ATM terminals, unaware of being part of an indoor radio network.

Figure 2.3 shows that the physical layer consists of the Physical Medium (PM) and the Transmission Convergence (TC) sublayers.

![Figure 2.3: The physical layer of the ATM reference model [de Prycker, 1991].](image-url)
The PM sublayer is responsible for the correct transmission and reception of bits on the physical medium. The physical medium function is medium dependent. The bit timing has to guarantee and regulate reconstruction at the receiver. This physical medium is studied in the indoor IBN project by Ir. P. Smulders and T. Wagemans.

The TC sublayer communicates with the PM sublayer using bit streams, and with the higher ATM layer in terms of streams of correct ATM cells. The cell rate decoupling function has to make the speed of the cell generation independent of the speed of the transmission network, as indicated by the word asynchronous in the acronym ATM. The Header Error Control (HEC) function calculates the error checksums of passing cells. When a cell is sent the HEC code is placed in the header, when a cell is received the HEC code is checked and the cell is rejected or accepted and passed. The cell delineation function has to transform an asynchronous cell stream into a synchronised cell stream.

One functionality below the cell delineation all the functions are independent of the functionality in the ATM and the higher layers of the protocol reference model. So this is a place where adaptation can take place.

The transmission frame generation/recovery receives bits from or sends bits to the PM sublayer. These bits form transmission frames, that can be of various types, e.g. SDH or cell based frames.

The transmission frame adaptation transforms these packets into the cells used in the cell delineation function.

The transmission problems in the adaptation of ATM to an indoor IBN should be dealt with in the transmission frame adaptation function of the physical layer.

2.6. Conclusions

The IBN in an indoor environment has been looked at in this chapter. The IRIBN has to become flexible, and compatible with the CCITT standardized ATM. ATM has as basic assumption that high quality links are used like optical fibers. Radio links don’t meet this requirement. This indicates that e.g. additional error control will be necessary. The adaptation of the radio network to standard ATM should be done mainly in the transmission frame adaptation function of the physical layer.
3. The Indoor ATM-based IBN

One of the conclusions of the previous chapter was that the indoor IBN should be an ATM network. In this chapter the indoor network is studied.

3.1. General description of the network

The Indoor Radio IBN (IRIBN) has to provide communication links between all the users in one room and with terminals outside that room. The outside terminals can be a network of rooms within one building or a public ATM network. These types of networks are not studied separately because they are beyond the scope of this study as mentioned in chapter 2. The outside world can be seen as being part of a standard ATM network. The connection point can either be seen as a network node of the outside ATM network or as an ATM-user network interface. This will depend on the size of the IRIBN.

The basic configuration has been studied by Sonnemans [1990]. He found that the I-IBN should consist out of a base station, e.g. in the ceiling of the room, and several terminals within reach of the base station. Connections can be established using the base station as switching point for all the traffic. It will be referred to as the central station.

ATM networks have to support three kinds of connections, which are:
1. terminal - terminal.
2. terminal - multi-terminal (multicast).
3. terminal - all terminals (broadcast).

All types of connections have to be possible, with terminals inside the room as well as terminals outside the room. The three types can easily be set up in the ATM layer. However, the hardware used in the physical layer has to serve all these different connection types.

3.2. Network Configuration

Several network configurations have the behaviour as described in section 2.1.. Two solutions assume a centralized hierarchical network. First, the pure star configuration can be used, which is illustrated in figure 3.1.

A radio link connects all the terminals with the central station in this configuration. All
terminals use the same frequency channel. Advantages of this configuration are that the number of links is minimal and that the transmitting and receiving equipment of all the terminals can be the same. Disadvantage is that only one frequency channel is used and that the configuration is thus relatively inflexible.

Second, the extended star network can be chosen, as illustrated in figure 3.2.

A group of terminals is connected with a concentrator in this configuration. The concentrator
is connected with the central station. This configuration has two appearances.

First, each concentrator is a physical independent entity standing alone in the room. Radio links or wires connect the central station to the terminals. The terminals connected to one concentrator use one frequency channel. Each concentrator uses its own frequency channel. Thus a multi channel environment or frequency multiplex has been created. This configuration more flexible than the star network. It becomes very easy to expand the capacity of the network, it is just adding one more concentrator. Another advantage is that within a very large room with many users, a concentrator can always be placed near the users. This makes is easy to fulfill the in line of sight requirement. Main disadvantage is that the number of links increases compare to the pure star configuration. If the concentrators and the central station are connected with wires, it is possible to locate them in different rooms. This can be advantageously for a small room next to a large room. However, these interroom networks will be no part of this study. Eventually, they can be subject of future studies.

Second, the central station and concentrator can form of one physical entity. The total number of communication links in the room is minimized. This number is the same as in case of a pure star network. The installation of remote concentrators is made superfluous. This reduces the installation costs. The capacity can be easily expanded by adding a new frequency (equivalent to one more concentrator). All the concentrators are within a line of sight of all the remote terminals. In theory, all the terminals can use all the frequency channels.

The extended star network configuration has two advantages over the pure star network. The available bandwidth is used more efficient if it is shared by all the users. Further, all the terminals can use the same transmitting and receiving equipment. This will probably lower the costs and more compatibility for the terminals. Uncertain is what the technical consequences are of this proposal. Questions are for example how many frequencies can be received and transmitted with one single antenna. This can be subject of further study.

The extended star network with the concentrators inside the central station with multi frequency channels seems to be the best configuration. This configuration will be used in the further part of this study. However, it is still uncertain or this is technical and economical a realistic solution. If not, it will be probably not to difficult to adapt the system to another configuration.
3.3. Access to the network

Most important of every network is that it can be used by all the connected terminal, i.e. the users have to be able to access the network and transmit their data. When a terminal wants to establish a connection in ATM, this is 'asked' to the ATM Call Acceptance mechanism. This mechanism can give a terminal access or deny the request.

ATM is a connection oriented protocol. This means that before the information transfer starts, a (virtual) connection has to be established. A connection is a route through the network which is travelled by all the cells belonging to that specific call. This is part of the signalling. ATM uses separate channels for this process. If no (physical) capacity is available for such a route, the access will be denied by the ATM call acceptance mechanism.

The correct operation of the call acceptance is crucial. If too many calls are accepted, the network will become overloaded. On the other way, if too many requests are denied, the system efficiency will become poor. These functions are performed above the physical layer. This indicates that the IRIBN, which operates in the physical layer (see chapter 2), has only to accept these decisions.

The reservation of bandwidth on the ATM layer level can be projected on the physical level. In the IRIBN the users ask the central station for physical capacity when they have cells ready for transmission (see chapter 2). A sequence of cell transmissions in the physical layer will be denoted as a session. This is done only after the ATM higher layers set up a connection. A access control scheme is necessary for this task. When bandwidth is given to a user, the data transmission can start. This bandwidth allocation is worked out in the next section.

**Definition:** A session is a burst of cells transmitted within an ATM call. An access control mechanism allocates the bandwidth for sessions in the physical layer.

Here arises a responsibility problem. In chapter 2 was concluded that the ATM layer shouldn't be aware of the medium which is used in the physical layer. It can be said that on one hand the physical layer is responsible for the data transport. On the other hand the ATM layer must decide to accept a call or not, depending on the available network resources. But the capacity available in the radio network depends on how the physical layer uses it, i.e. the physical layer has its own allocation mechanism. The allocation mechanisms of the physical and the ATM layer have to work together in a decent way.

When a terminal wants to start cell transmission in an ATM channel (i.e. a session), it has to request the access control functionality for capacity. A control channel is used by all the
users for this purpose. All the terminals should have the same chances to access this channel. To regulate the traffic on the control channel it is necessary to have a fair and correct multi-access protocol. This protocol avoids that messages of the terminals collide and destroy each other. When a terminal uses this protocol, it can get access to the control channel. When access is gained, the control information can be send by the terminal. The control information is processed in the central station which can allocate a specific part of an information channel to the requesting terminal. The control and information channels in the physical layer are drawn in figure 3.3.

![Figure 3-3: Structure of the subbands of the total available bandwidth of the IBN.](image)

3.4. Channel allocation

The central station knows which terminal requests for starting a session. Remains as problem on which channel the terminal is going to be permitted to send its data. It is assumed that all the terminals can use all the frequency channels (see section 3.2.). This is a typical bandwidth allocation problem, i.e. how the bandwidth has to be divided. The four possible solutions can be seen in Table 1.

First, the channel bandwidth can be taken variable or constant. A fixed bandwidth indicates a fixed bit rate per channel. This will be most simple to realise: the central station and the terminals can use always the same oscillators. A variable bandwidth indicates that in case of low system load a low bit rate can be used. It is almost evident that the 'propagation' performance (e.g. BER) will than increase. Disadvantage is that the transmission delay will then become higher and the transmitters and receivers more complex. In case of heavy loads
Table I: Bandwidth allocation mechanisms.

<table>
<thead>
<tr>
<th></th>
<th>channel bandwidth</th>
<th>channel bandwidth</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>fixed</td>
<td>fixed</td>
</tr>
<tr>
<td>fixed terminals</td>
<td>type 1</td>
<td>type 2</td>
</tr>
<tr>
<td>per channel</td>
<td></td>
<td></td>
</tr>
<tr>
<td>variable terminals</td>
<td>type 3</td>
<td>type 4</td>
</tr>
<tr>
<td>per channel</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A higher bit rate can be used. It seems that the first option is the most realistic one in the current technical situation.

Second, which terminals are permitted to use which channel? From a traffic point of view, it is most advantageously if all the users can access all the channels. This increase the complexity of the system (see for details section 2.2.). While ATM has to become flexible, all users should be permitted to use all channels. This will of course increase the complexity of the bandwidth allocation algorithm. The division of the active users over the channels in the best way, needs a sophisticated algorithm.

To reduce complexity, in the next part of this study is assumed that each connection consists of only one fixed channel, which is allocated in the "call set up" phase. From channel can be switched eventual advantageously during the call (rearranging).

A good theoretical analysis or simulation can be done to search for the best allocation method. Choosing criteria can be:

1. Complexity.
2. Technical possible.
3. Flexibility.
5. Throughput.

Complexity implies almost always difficult to realise. The complexity has to be kept minimal because the receivers and transmitters are probably the most expensive parts of the network, so their price is very important for the success of the network.
3.5. Conclusions

The IRIBN should be an extended star network with the concentrators in one physical entity with the central station. The bandwidth is divided in control and information channels. A multi access protocol is needed to perform the access to the control channel. A session is defined as a burst of cells transmitted within an ATM call. ATM cells are only transmitted within a burst. An access control scheme in the physical layer performs the bandwidth allocation for sessions. Another bandwidth allocation mechanism decides how the users are spread over the information channels.
4. The multi access protocol

4.1. Tasks of the protocol

A multi access scheme is needed to regulate the access to the control channel(s) of the physical layer (chapter 3). Some tasks of the multi access protocols were mentioned in chapter 2.3. The choice of a particular protocol that fits these requirements is studied in this chapter.

The transport of data between the central station and a terminal, as desired by the ATM-layer, is a major task. The connections are requested using a common control channel. The terminals, which are ready for transmission, compete with each other for this channel. This competition has to be effective and as fair as possible. This to reach acceptable results. The total delay caused by this protocol is very substantial and has thus to be kept minimal.

It has to be noted that the multi access protocol is only used within established ATM connection, to start a session.

4.2. The protocols which can be used

Many multi access protocols are known. A brief survey of them is given in [Tobagi, 1990]. Every protocol has it's advantages and disadvantages compared to the others, e.g. in complexity, throughput or delay. A major characteristic of each protocol is the randomness of the access. The extremes are fixed assignment and totally random. Fixed assignment means that every terminal is only permitted to transmit during a specific time (Time Division Multi Access (TDMA)) or in a specific frequency (Frequency Division Multi Access (FDMA)). In a random protocol each terminal can transmit its messages whenever it wants, like the ALOHA protocol. The ALOHA protocol was introduced in the early seventies by the University of Hawaii.

In general, fixed assignment techniques are most advantageous in case of heavy loaded networks. This because the whole available bandwidth can be used if all the terminals fill their assigned part of it. Purely random protocols are preferred in networks with a large number of users with, compared to the network capacity, small traffic which is bursty. All the available multi access protocols are, concerning there randomness, lying somewhere between these extremes.

A major restriction for the protocol which is going to be used in the IRIBN is that the remote terminals will be unable to hear each other. With this handicap, only a few of the known protocol are useful. This problem is known as the hidden terminal problem.
The division of the bandwidth (as suggested in the previous chapter), in a multi access, a control information and N information channels, is related to the Split-Channel Reservation Multi Access (SRMA) type [Tobagi & Kleinrock, 1976]. With hidden terminals, it can be used in combination with Busy Tone Multi Access (BTMA) or Idle Casting Multi Access (ICMA). A totally different solution is polling. In the next sections these protocols will be described and analyzed.

4.3. SRMA

SRMA is a family of protocols. It is described in [Pach, 1989] and was introduced by [Tobagi & Kleinrock, 1976]. It has been proposed as alternative for the conventional polling schemes for communication between a set of users and a central station. SRMA uses an explicit reservation technique—the user who has a packet to transmit first makes the request for service on some subchannel. The central station manages the queue of requests and informs the user about his granted time frame for transmission of the packet.

4.3.1. Operation of SRMA

The available bandwidth is divided in control (request and answer) channels and message (information) channels. These can be frequency or time multiplexed, depending on the system. See figure 4.1.

![Figure 4.1: SRMA channels.](image)

If a terminal has a packet ready for transmission, it makes a service request on the request channel, and waits for an answer from the central station. The central station manages the FIFO queue of requests and informs the users when they are permitted to transmit their
messages (session is started). Because the service request has to contain a source address, the acknowledgement can be a relayed version of this request. If the waiting time exceeds the Time Out time ($TO$), the terminal generates a kind of interrupt, and a new service request is made. Duplicate requests in the FIFO queue are prevented by the central station. All known SRMA schemes in the literature, use only one message (information) channel. This makes it possible to use one channel for both the acknowledgements and the messages. This is called the Request Message (RM) scheme. Separate answer and message channels are used in the Request Answer Message (RAM) scheme. This indicates simultaneous acknowledgement and message transmission. This makes the system less sensitive for propagation delays between the terminals and the central station.

In a multiple message channel environment the RM scheme will be difficult to realise, because a terminal doesn’t know on which message channel it can expect the answer. This makes simultaneously receiving of all message channels necessary, what probably will be very expensive and complex. Solution can be to use only one message channel to transfer acknowledgements. This results in a system comparable with the RAM scheme. This indicates that only the RAM protocol will be useful in the IRIBN with a multiple message channel configuration.

The request channel is in general used in a random access mode like CSMA. In this particular case another protocol, will be necessary.

4.3.2. Analysis of SRMA

SRMA includes a lot of variables, as shown in the previous section. To reach best performance an optimal combination has to be searched for.

It is interesting to look at the delay introduced by SRMA. To do this in a proper way the following variables will be used [Tobagi & Kleinrock, 1976]:

- $b_r$: length of the request packets.
- $b_m$: length of the message packets.
- $b_a$: length of the answer packets.
- $\eta_a = \frac{b_a}{b_m}$: ratio of the length of the answer and the message packets.
- $\eta_r = \frac{b_r}{b_m}$: ratio of the length of the request and the message packets.
- $\eta = b_a = b_r$: total available bandwidth.
- $W_r$: bandwidth of the request channel.
- $W_m$: bandwidth of the message (information) channel.
- $W_a$: bandwidth of the answer channel.
The total SRMA delay can be split in two parts. Delay 1 \((D_1)\) is the time necessary for a request to be successful received at the central station. Delay 2 \((D_2)\) is the time between reception of the request packet and the end of the message transmission. These delays are visualized in figure 4.2. The delay caused by the SRMA protocol can be seen as the session delay, i.e. time needed to start a session.

![Figure 4.2: Packet delay in SRMA-RAM protocol.](image)

### 4.3.2.1. Delay \(D_1\)

Figure 4.2. shows that \(D_1\) is independent of the information channel allocation mechanism as described in chapter 3. It will depend only on the request mechanism used in SRMA. The delay \(D_1\) consists of four components.

First component is the time needed to transmit a request \((T_r)\). This delay depends upon the length \(b\), and the available bandwidth \(W\), or the bit rate of the request channel (the bit rate is the same in bits as the bandwidth in Hz. if a modulation of 1 bit/Hz. is used, which is
assumed in the rest of this chapter). This component is equal to: $\eta/(1-\theta)$. The transmission delay becomes very small when very high bit rates are used.

The second component is caused by the number of retransmissions ($N_r$) necessary to receive an acknowledgement. Each retransmission implies one Time-Out ($T_{to}$ s) and a new request $\eta/(1-\theta)$.

Third, the (random) (re)transmission delay ($D_r$) is taken into account. This is the time the access protocol needs to generate and schedule a new request in case of a Time Out. This has to be a random time, otherwise multiple collisions will occur after a request collision, and a dead lock loop will be entered.

Fourth all signals have to travel from the terminal to the central station or vice versa. This results in a propagation and delay time ($\tau$). With short distances this delay will become very small. A high bit rate is for relatively narrowband channels only possible if it is time multiplexed with other signals. A time multiplexed SRMA may be preferable above a frequency multiplexed system.

The four components give for the delay $D_1$ [Pach, 1989]

$$D_1 = \tau + \frac{\eta}{1-\theta} + (N_r - 1) \left( \frac{\eta}{1-\theta} + T_{to} + D_x \right).$$

This equation can be transformed in a different way. This is done by splitting the variables in a category influenced by the multi access protocol on the request channel and a category introduced by the SRMA protocol. This has been done in [Tobagi & Kleinrock, 1976], which gives

$$D_1 = D_{req.protocol}(S_r) \cdot \frac{2\eta}{1-\theta} \left[ \frac{b_m}{W} \right].$$

So the parameters influencing the SRMA scheme and which are independent of the request multi access protocol are $W$, $b_m$, $\theta$ and $\eta$.

### 4.3.2.2. Delay $D_2$

The delay $D_2$ consists of some components. The transmission and the propagation delays are of the same type as in section 4.3.2.1. Without these, only the request ($D_{rq}$) and the packet queuing ($D_{pq}$) time remain.

The time $D_{rq}$ is needed for the assembly and the queuing of an acknowledgement in the
central station. In [Tobagi & Kleinrock 1976] is derived that if the inequality

\[ W_a \geq W_r \frac{b_a}{b_r} \]  

(4.3)
is satisfied no queuing delays will occur in the central station. This means that with a given request channel bandwidth and the size of the request packets, the ratio of the answer channel bandwidth and the size of the acknowledgements can be calculated.

After the terminal has been acknowledged, it has to wait a time \( D_{pq} \), as has been calculated by the central station. This time is announced in the acknowledgement. It is obvious that this delay isn’t influenced by the multi access protocol of the request channel. It can be calculated only after the channel and bandwidth allocation mechanism is chosen.

So the parameters which are of concern in this section are \( W_r, W_a, b_a \) and \( b_r \).

4.4. Request channel multi access protocols

In this section the alternatives for the multi access protocol of the request channel of SRMA-RAM will be studied. First will be described what the contents of a request packet should consist of.

The request channel multi access protocol manages the admittance to the request channel of SRMA. Before the alternatives will be discussed, the size of the requests is estimated. The contents of a request has to consist of (at least):

1. Local address of the requesting terminal. As maximum number of terminals is taken 200 [Sonnemans, 1990]. To discriminate them at least 8 bits addresses are necessary.

2. Type of connection requested for. Until now 4 types of traffic are expected in ATM. The channel allocation mechanism in the central station probably can use this type to use it for the traffic handling. For this purpose are at least 2 bits necessary.

3. The number of packets ready for transmission. With this number, the central station schedules the traffic on the information channels. In case of continuous traffic this field contains the bit rate of the sending terminal. The bit rates will be up to 150 Mbit/s, which is a 28 bits number (assuming a step size of 1 bit).

4. It should include a header error check, to assure correct reception of the requests. 8 bits are taken for this purpose [Sonnemans, 1990].

5. A preamble is necessary for synchronization of the central station. This
schemes. [R. Sonnemans, 1990] uses 8 bits for this purpose. 

So the size can be approximated to be minimal 56 bits.

4.4.1. Busy Channel Multi Access (BCMA) 

Busy Channel Multi Access has been introduced by [Andrisano, 1990]. In this protocol the available bandwidth is divided in two channels, a busy channel and a transmission channel. If necessary, it is possible to use a couple of transmission channels using only one busy channel. However, in this section only the one transmission channel configuration will be studied. This is illustrated in figure 4.3.

![Figure 4.3: BCMA sub channels.](image)

The busy channel is managed by the central station and carries information about the state of the transmission channel. This will be the busy or the idle state. The central station generates and transmits the busy signal. All the terminals receive this signal. When a user wants to start transmission, it senses the busy channel. If the channel is assumed to be in the busy state, the terminal aborts and reschedules its request. If the busy channel is sensed free, the terminal transmits its request. Else, it reschedules its attempt. If wanted, the busy channel can transmit SRMA acknowledgements instead of the signal belonging to one state.
4.4.1.1. Performance of BCMA

In this section the performance of BCMA will be looked at. Especially the variables influencing a successful transmission of a request will be mapped.

The probability that a packet will be transmitted successful through the network ($P_s$) is equal to the probability a terminal is permitted to send ($P_{ps}$) multiplied by the probability of correct transmission ($P_{ct}$), which gives

$$P_{st} (m) = P_{ps} (m) \cdot P_{ct} (m) \quad .$$  \hfill (4.4)

The variable $m$ means the $m$-th terminal. This variable is introduced because the probabilities can differ for each terminal. However, from now on they will obeyed, which indicates that is calculated with mean values.

The probability $P_{ps}$ depends on both the inside ($P_{ips}$, due to the retransmission scheme after an unsuccessful access attempt) and the outside ($P_{ops}$, strictly correlated to the state of the busy channel) permission to send probability, thus

$$P_{ps} = P_{ips} \cdot P_{ops} \quad .$$  \hfill (4.5)

The probability $P_{ips}$ will be a design parameter. It is correlated with the random retransmission delay $D_r$ (section 3.3.2.1). The outside probability will be

$$P_{ops} = P_{fc} \cdot P_{ct/fe} + P_{bc} \cdot P_{wi/bc} \quad ,$$  \hfill (4.6)

with for the subscripts

- $ci$ correct investigation of the busy channel,
- $wi$ wrong investigation of the busy channel,
- $bc$ 'busy' on busy channel,
- $fc$ 'free' on busy channel.

The probability $P_{ct}$ of correct transmission depends on both the characteristics of the transmission channel ($P_{cic}$) and the probability, when transmission is permitted, no collisions occur ($P_{nc}$). This gives

$$P_{ct} = P_{nc} \cdot P_{cic} \quad .$$  \hfill (4.7)

The characteristics of the information channel will not be influenced by BCMA. If the channel can be assumed to be ideal, $P_{cic}$ is unity.
The no collision probability $P_{nc}$ is the complement of the collision probability $P_c$. It is obvious that the performance of BCMA degrades with the number of collisions on the transmission channel. Collisions have two main causes.

When a terminal starts transmission, the busy channel moves to the busy state. This takes of course a certain amount of time. This will be at least the propagation time $\tau$ between terminal and central station. When the central station receives the request, it takes some time to detect it (due to synchronisation), resulting in a packet detection delay time $(D_p)$. After the channel has been put in the busy state, it takes another time $\tau$ for the other terminals to receive this busy signal. This means that they assume the transmission channel to be free a time $2\tau+D_p$. In this period collisions can take place although everything operates in a correct way. Another reason for collisions is incorrect interpretation of the transmitted busy channel signal. This is due to a wrong detection in the receiver.

4.4.2. Busy Tone Multi Access (BTMA)

BTMA is a special case of BCMA, which has been described in the previous section. However, BTMA was introduced before BCMA [Tobagi & Kleinrock, 1975]. This protocol can be chosen as multi access protocol for use in the IRIBN [Sonnemans, 1990]. This was taken without being concerned as part of a SRMA protocol.

In BTMA, the busy channel is a tone, which is turned on and off. It is showed that the performance can be compared with standard Carrier Sensed Multi Access (CSMA) with a doubled transmission delay. CSMA is a protocol familiar to ALOHA. Big difference is that a terminal senses the message channel before starting with the transmission of a packet. If another terminal is detected as active, the packet is rescheduled.

In figure 4.4. [Andrisano, 1990] the throughput of BTMA is compared with ALOHA (curve a) and CSMA (curve b). $t_{da}$ is the time the busy channel is sensed (normalized) normalised to the message packet size (in fact a SRMA request packet!) The time $t_{da}$ is strongly influenced by the choice of the receiver. In the BCMA analysis this time is hidden in the variables $P_{ops}$ and $P_c$. These are time dependent because of the propagation and packet detection time $2\tau+D_p$. When the channel is sensed for a long time, the effect of these delays becomes of less significance.
From figure 4.4. can be seen that BTMA with a very short busy channel detection time $t_{dn}$ has about the same throughput as CSMA. CSMA is nowadays a very frequent used multi access protocol in LANs, and is in general seen as a protocol with a good performance. However, when the detection time becomes of the same size as the message packets of BTMA (SRMA request packets!!) the performance degrades to the ALOHA level.

In the literature the performance and delay studies are performed for analogue receivers, with the aim to find optimized solutions and the corresponding parameters for this case. In an integrated digital network this doesn't seem ideal. In the next section Idle Casting Multi Access (ICMA) will be discussed. This is in principal a kind of improved BTMA with a lot of possible additive features.

4.4.3. Idle Casting Multi Access (ICMA)

In the late seventies the (Japanese) NTT introduced the ICMA-family of multi access protocols in their land mobile telephone system. In ICMA, the BCMA busy channel is called the idle channel. The protocol behaves besides this exactly the same as BTMA. The multiple seizure or collision probability $P_c$ of ICMA can be seen in figure 4.5 [Okasaka, 1978].
Figure 4.5: Multiple seizure probability.

The time $t_l$ of figure 4.5 corresponds with the packet detection delay $D_p$ from section 4.4.1.1. It can be seen that with an increase of this time the multiple seizure probability increases too. This was already predicted in section 4.4.1.1.

4.4.3.1. ICMA-CD

With the aim to reduce the number of collisions, ICMA Collision Detect (ICMA-CD) was suggested [Murase, 1987]. This was done analogous to the collision detect variant of CSMA. In this variant the sending terminals try to detect or their transmitted messages collide. If a collision is detected, the message packet will be rescheduled. Advantage of such scheme is that the terminals don't need the central station or other terminals to detect a collision, which reduces the waiting times in the system and makes a quick reaction of the terminals possible (abort the transmission). Thus the message channel becomes earlier available for a new transmission attempt. This can be seen as a reduction of the number of collisions while only a few packets of a message will collide.

In ICMA-CD the central station verifies the 'error checks' from the words (slots) of the received messages (again SRMA requests). If the error check is incorrect, the central station assumes a collision has taken place. In this case the central station transmits a STOP signal.
on the idle channel and the all terminals abort their transmission. An example of the behaviour is given in figure 4.6. (a) is an ICMA and ICMA-CD successful transmission. It can be seen that in case of successful transmission both protocols have exactly the same behaviour. In (b) is given an example of a collision in ICMA. The collision continues until the whole messages are transmitted. This means that for the whole message period the downward channel will be occupied. In (c) is given an example of a collision in the ICMA-CD case. After having received the first word of the messages, the central station concludes there has been a collision, and transmits a STOP signal. Now the downward channel becomes already available after the transmission of the second word of the messages.

With the preamble signal the central station detects the start of a packet transmission and synchronizes the bit timing (see 4.4).

The improvement in performance is thus gained with a shorter collision recovery time after a collision (enter the idle state again) compared to ICMA. In figure 4.7. some simulation results are presented [Murase, 1987].

Figure 4.6: ICMA en ICMA-CD example.
This picture shows again that the packet detection time $D_p$ influences still the performance. Further, and what is of more concern in this section, can be seen that ICMA-CD performs better compared to ICMA if the throughput is concerned. The throughput performance of ICMA is like BTMA somewhere between Aloha and CSMA.

In [Adachi, 1988] is shown that this non-persistent (random retransmission time after a collision) protocol performs better than the 1-persistent protocol (always immediate retransmission after a collision) variant. Hence the non-persistent protocol variant is preferably.

4.4.3.2. Improved ICMA-CD (I-ICMA-CD)

Another feature can be added [Choi, 1989]. The channels are assumed to be slotted now. A terminal is only permitted to start transmission at the beginning of a slot. This can be seen as a protocol addition like the slotted version of the ALOHA protocol. After a terminal starts transmission, it counts the number of slots it takes for the idle channel to become busy. If the corresponding time is shorter then the packet detection delay, the terminal initiated his transmission after another terminal did, and aborts transmission. If the times are equal, the message didn't cause a collision and continues. However, when two terminals start simultaneously, neither of them aborts and a stop signal will still be necessary.

To avoid disruption of the messages, a preamble longer then the packet detection delay has to be added. In the proposal of [Choi,1989] the preamble necessary for synchronization and bit timing is omitted. This implies a perfectly slotted system. This depends on the quality of the receivers and transmitters, and has to be further studied.

The above described improved protocol performs better then the ICMA-CD protocol.
[Choi, 1989]. To achieve this gain, system complexity will not increase very much.

4.4.3.3. Recent ICMA-variants

ICMA-DR

The main shortage of ICMA is the difficulty to shorten the collision recovery time, because terminals can’t recognize that their packet transmission has collided until the information is received by the terminal from the central station.

To improve this [Murase, 1989] suggested the Data Slot Reservation (ICMA-DR) ICMA variant. When the transmission/upward channel is sensed idle, the terminals start sending a request on the upward channel (Data Slot Reservation DSR packet). This is a very short packet. After the reception the central station can give a Data Packet Request (DPR) and the message transmission can start. Murase shows the performance is superior over standard ICMA.

This is a kind of SRMA-Request protocol. It offers most gain in case of relatively large messages. The function of the multi access protocol in the IRIBN is to transfer relatively short request messages. A much more complicated protocol is the consequence.

VT-ICMA

The last variant of the ICMA family are the virtual time protocols. These protocols use two times: the arrival and the transmission time. They can be used in combination with the above described collision detect and reservation variants. They seem to increase the performance. However, they are again more complex and are out of the scope of this study. Interesting references are: [Lee, 1989], [Song, 1989] and [Molle, 1985].

4.4.4. Polling

Polling is a relative old and simple multi access protocol. Every terminal is periodic ’pollled’, irrespective or it has packets ready for transmission or not. Polling is thus a kind of a fixed bandwidth allocation mechanism. This is most advantageous in heavy loaded systems, in contrary to random access, as mentioned in 3.2. Token ring systems, used in LANs, are distributed polling systems.

In the system configuration chosen in the previous chapter, only centrally controlled systems are useful. The roll-call polling technique matches this requirement [Tobagi, 1976]. In this
protocol all the terminals are sequential polled. If a terminal is polled, it empties the transmission buffers. This implies (with the assumption of infinite buffer length) that all packets which are queued can be transferred. This makes a maximum channel utilisation of one possible. The disadvantage is that the queuing delay, which are of key interest in ATM, can become very large in heavy loaded systems, especially in the case of many users. This is illustrated in figure 4.8.

![Figure 4.8: Packet delay in roll-call polling (a=0.01).](image)

**Explanation of figure 4.8:**

- $M$: Number of terminals.
- $L$: Ratio of the size of the message packets and the polling packets.
- $a$: Ratio of the propagation time, between central station and terminal, and the transmission time of a message packet.

From figure 4.8, can be concluded that using polling is most advantageous in an environment with a relatively small population ($M$) and a lot of traffic ($S$).
The roll-call polling technique uses only one channel for both the polling and the message (information) packets. In the IRIBN was decided to use separate channels. The number of users is estimated to be around 200. In figure 4.8, can be seen that \( M = 100 \) already implies very large delays. Because the message and the polling packets will have approximately the same size, \( L = 1 \), the performance will drop to an unacceptable level. Hence polling isn’t a good alternative for SRMA.

4.4.5. Relations between SRMA and Polling

Polling and SRMA can be seen as totally different protocols, they have also a lot of things in common. In classical polling systems, as described in section 3.4, a user is polled and is then permitted to start transmission. In SRMA a terminal 'polls' the central station and receives an acknowledgement. So SRMA can be seen as a kind of reverse of 'selective' polling. With a lot of inactive terminals this being 'selective' is of course most advantageous.

Besides this, it may be advantageous to avoid that users of continuous services, like video or telephone, use the SRMA request channel protocol every time they want to transmit a packet. When they are in a session, they don’t need the request channel anymore. In such a session they can be polled. This reduces the traffic on the request channel, because there are less contending users. That reduces the delays as introduced by the request channel. This will be subject of study in chapter 6.
4.5. Conclusions

The described system in chapter 2 can be seen as a SRMA protocol, with N message (information) channels. The SRMA-RAM variant is most suitable for being used in the IRIBN. A time multiplexed system may be preferable above a frequency multiplexed system. A BCMA protocol can operate the request channel. The packet detection time in the central station influences especially the performance.

The BTMA and ICMA protocols were looked at as being a BCMA implementation. The improved ICMA-CD protocols is the best choice.

It may be advantageous to use the polling technique to reduce the traffic on the SRMA request channel. This can be done for continuous bit rate connections as telephone and video. They can be polled within a session using the SRMA answer channel.

It hasn’t been possible to calculate the throughput and delay of this proposal/choice. This has a couple of reasons. It will be first necessary to estimate the traffic characteristics of the terminals. Second, it is necessary to know the delays caused by the bit timing and synchronisation on the link level. Third, the quality and the bit rate of the transmission links are not known.
5. The Improved ICMA Collision Detect Protocol

In this chapter the I-ICMA-CD protocol, as described in section 3.3.2.3., will be analyzed [Choi, 1989]. The aim is to derive the average packet delay and the throughput as introduced by this request channel protocol for the indoor 60 GHz channel.

5.1. Protocol description and assumptions

The I-ICMA-CD protocol is slotted, i.e. the time axis is divided in several small periods of constant length, named slots. Starting a transmission is only permitted at the beginning of a slot, which increases throughput performance (see figure 4.7 for the ALOHA case). The establishment of a slotted network out of separated radio links, includes many difficulties. However, in this chapter is assumed that a slotted network is available.

It is obvious to take the length of the idle/busy signals (8 bits [Okasaka, 1978]) as the slot size $A$, as shown in figure 5.1.

![Figure 5.1: Time Slot of the network.](image)

With a bandwidth of $W_c$ MHz, the transmission of 8 bits takes $8/W_c \mu s$ (assuming a 1 bit/Hz. modulation, see chapter 4). Further, a guard time is necessary, to make synchronisation possible, which is equal to the propagation delay time $\tau$. With a maximum link length of 30 m [Sonnemans, 1990], $c=3\times10^8 \text{ m/s}$, $\tau$ equals 0.1 $\mu$s. The slot size $A$ is given by

$$A = \left(\frac{8}{W_c} + 0.1\right) \mu s. \quad (5.1)$$

The collision duration $B$ ($D_p + 2\tau$, determining $P_n, in \ formula 4.7, section 4.4.1.$) can be expressed in a number of slots (with size $A$). However, slotted systems use only discrete time intervals. To overcome this problem, the number of slots is rounded off to the nearest larger integer. The size becomes then $Y (\geq 2)$ slots, and can be expressed by

$$\lceil B \rceil = A \cdot Y \mu s. \quad (5.2)$$

In section 4.4. is derived that the request message size has to be at least 56 bits. Here a
message size of 128 bits is taken, including an additive 8 bit error check to make the CD scheme possible. The preamble has a duration of the collision period $B$ in slots, plus one extra slot for the transmitting and the reception of a busy signal (see section 5.2.). With this preamble of $Y+1$ slots, the total request message length will become $Y+17$ slots. For $Y=2$, this is shown in Fig. 5.2.

| P | P | P | 1 | 2 | 3 | 4 | h | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |

Preamble (with $Y=2$!), Message slots 1 t/m 15, h = error checksum.

Figure 5.2: ICMA-CD Request Message.

A Bit Error Rate (BER, $P_e$) of $10^{-6}$ is taken for all bandwidth between 0 and 150 Mbit/s. However, in the sections 5.2. & 5.3., the channel is assumed to be ideal, indicating $P_e=0$.

5.2. The protocol scheme

When a terminal wants to start the transmission of a message, it will sense the idle channel. If it is assumed to be in the idle state, the terminal starts the transmission of the preamble at the beginning of the next time slot (see figure 5.3).

**IDLE CHANNEL**

```
I I I B B B B B B B B B I I I
```

$I =$ idle, $B =$ busy.

**SUCCESSFUL ATTEMPT ON REQUEST CHANNEL**

```
P P P M M M M M M M M M M M M - -
```

$P =$ preamble, $M =$ message, $H =$ error checksum. Slot counter: 3.

**UNSUCCESSFUL ATTEMPT DETECTED BY SLOT COUNTER**

```
P P -
```

$=$ no more transmission, random rescheduling. Slot counter $\leq 3$.

Figure 5.3: Successful request message transmission for $Y=2$.

A slot counter, which counts the number of time slots going by, is immediately started. After $Y$ slots the central station detects the transmission of the remote and the next slot the idle channel is moved to the busy state. When this busy signal is received by the initiating terminal, i.e. its slot counter equals $Y+1$ (in this example 3!), it continues the transmission of the request message. A terminal which initiated transmission in one of these $Y+1$ slots, has, when it receives the first busy signal, a slot counter (value) $\leq Y$ and stops the
transmission. If there wouldn't have been counter, this terminal would have been unable to determine that the busy signal wasn't the response to its own transmission.

The situation in which two users try to get access to the request channel simultaneously, is shown in figure 5.4. In that case two terminals sensed the idle channel idle in the previous slot.

**IDLE CHANNEL**

| I | I | I | B | B | B | B | B | S | I | I | I | I | I | I | I | I | I |

B=busy, I=idle, S=stop signal after collision detection.

**USER 1**

| P | P | P | M | M | M | M | H | M | - |

P=preamble, M=message, -=abort. Slot counter = 3.

**USER 2**

| P | P | P | M | M | M | M | H | M | - |

Slot counter = 3.

Figure 5.4: Collision detection by CD Scheme, for Y=2.

While two (or more!) accessing users started their message transmission at the beginning of the same slot, the contents of their slot counters will be equal. This indicates that when the first busy signal is received, both terminals will continue transmission, and the messages collide. The central station can detect such a collision by receiving an incorrect error checksum (fifth slot!). At the next time slot the idle channel is then moved to the stop state, and after the reception of this signal both users will abort their transmissions.

It is possible to move the idle channel to the idle state one slot before the end of a request message, because the requests have a constant length. During the stop signal the non transmitting users can behave the same as in case of an idle slot, so that the next slot transmission can started. This can be seen in figure 5.3 and 5.4.

The above described behaviour can be translated in a state description. For the state diagram see figure 5.5.

When the idle channel is in the idle state and none of the terminals has yet started transmission, the protocol is in the Empty State (ES). The ES has a duration of one time slot. The number of terminals, initiating transmission during this time slot (i.e. start of transmission at the next time slot), determine the next state. When there is no attempt, the next state will be the ES again. With one attempting user the Success State (SS) is entered.
Figure 5.5: I-ICMA-CD protocol state diagram.

(see figure 5.3). Finally, when there are two or more competing users, the Collision State (CS) is entered (see figure 5.4).

The Protocol Cycle is defined as a chain of protocol states, beginning at a ES and ending just before the next ES. From figure 5.5 can be seen that there are three protocol cycles possible: ES, ES-SS and ES-CS. The overall protocol behaviour is no more than a sequence of these protocol cycles.

In the further analysis will be assumed that the probability of one and only one request attempt in a slot, $P_r$, is equal for each slot. This assumption is only valid if some conditions are fulfilled. First, the population must have such a size that individual terminals don’t determine the overall traffic. Second, all the users have to be independent of each other. Third, the rescheduling scheme has to be random.

So $P_r$ depends on the traffic characteristics of all the users together and the slot time $A$. The slot time is determined by the bandwidth of the request channel (formula 5.1).

5.3. Protocol Analysis

In this section the performance of the I-ICMA-CD protocol is analyzed. However, this is done in a different way from section 4.4.1. for BCMA. The model of BCMA gives a good insight which parameters have influence on the performance. In practice it is difficult to follow this abstract model. The state transition probabilities out of figure 5.5 will be calculated. The probability $P_a$ that exactly one request appears in the ES slot, is given by

$$P_a = P_r.$$ (5.3)
The probability of two or more arrivals in the ES slot \( P_b \) can be expressed as

\[
P_b = P_r^2 + P_r^3 + P_r^4 + P_r^5 + \ldots
\]

\[
= P_r^2 \left( 1 + P_r + P_r^2 + \ldots \right) = \frac{P_r^2}{1 - P_r}.
\]

Finally, the probability of no request attempts \( P_{no} \) is given by

\[
P_{no} = 1 - P_a - P_b = 1 - P_r - P_r^2 - P_r^3 - \ldots
\]

\[
= 2 - \sum_{n=0}^{\infty} P_r^n = \frac{1 - 2P_r}{1 - P_r}.
\]

To approximate the average number of reschedulings and the average delay, a closer look to the protocol cycle is necessary. First off all the mathematical expectation \( E\{CT\} \) of the length of the protocol cycle will be calculated. Each cycle consists of at least one idle slot (ES). The duration of the SS will be (on the idle channel) \( Y+16 \) slots, \( Y \) idle states and 16 busy states. This occurs with probability \( P_a \). The duration of the CS is only \( Y+6 \) slots, \( Y \) idle and 6 busy states. Together this gives an \( E\{CT\} \) of

\[
E\{CT\} = 1 + P_a (Y+16) + P_b (Y+6)
\]

(5.6)

Slots. Combining formulas 5.3-5.6 gives for the amount of idle states

\[
P_r (Y-1) + 1
\]

\[
\frac{1 - P_r}{1 - P_r}
\]

(5.7)

slots, for the amount of busy states

\[
16P_r - 10P_r^2
\]

\[
\frac{1}{1 - P_r}
\]

(5.8)

slots and finally for \( E\{CT\} \)

\[
E\{CT\} = \frac{P_r (Y+15) - 10P_r^2 + 1}{1 - P_r}
\]

(5.9)

slots.

The start of the message transmission can only start immediately after the ES. This is only one out of \( E\{CT\} \) slots. So \( (E\{CT\} - 1)/E\{CT\} \) of the request attempts has to be at least rescheduled. From the remaining \( 1/E\{CT\} \) requests only a part \( P_a \) of the cases will lead to
a successful transmission. The probability that rescheduling isn’t necessary \( P_{st} \) (see chapter 4, with no transmission errors) is

\[
P_{st} = P_s \cdot \frac{1}{E\{CT\}} = \frac{P_r(1-P_r)}{P_r(Y+15) - 10P_r^2 + 1} \quad (5.10)
\]

With this probability the average number of reschedulings \( N_r \) can be determined. This can be seen as follows. When a terminal makes a transmission attempt, it has always a probability \( P_{st} \) to make a successful request. If not, a rescheduling is made, and the same procedure is followed again. With \( P_{nst} = 1 - P_{st} \) this gives

\[
N_r = P(w=0) \cdot 0 + P(w=1) \cdot 1 + P(w=2) \cdot 2 + P(w=3) \cdot 3 + \ldots
\]

\[
= P_{st} \sum_{n=1}^{\infty} n \cdot (P_{nst})^n = \frac{(2-P_{nst}) \cdot P_{nst}}{(1-P_{nst})} = \frac{1-(P_{st})^2}{P_{st}} \quad (5.11)
\]

With: \( P(w=0) = P_{st} = 1 - P_{nst}, P(w=1) = P_{nst} \cdot P_{st}, P(w=2) = P_{nst} \cdot P_{nst} \cdot P_{st}, P(w=n) = P_{nst}^n \cdot P_{st} \).

Together with the random retransmission delay \( (D_r, \text{ see chapter 4}) \) and formula 5.11 the average delay can be easily calculated.

The throughput \( (S) \) can be calculated with formula 5.3 and 5.6. The ratio of the number of slots used for successful transmission (on the idle channel) and the total number of slots (on the idle channel) is taken for the throughput \( S \). This is done using the average \( E\{CT\} \), as given by

\[
S = \frac{P_r(Y+16)}{E\{CT\}} = \frac{P_r(Y+16)(1-P_r)}{P_r(Y+15) - 10P_r^2 + 1} \quad (5.12)
\]

It has to be noticed that \( S \) is that part of the channels that is used for the request transmission.

With formula 5.11 and 5.12 the throughput \( S \) and the average number of reschedulings \( W \) are expressed as function of \( P_r \).

5.4. Protocol performance with transmission errors

In the sections 5.2. and 5.3. is the performance of the I-ICMA-CD protocol calculated assuming there are no bit errors due to bad transmission. However, the BER \( P_e \) of the transmission channel will be about \( 10^{-6} \).
It has to be noticed that requests always have to be received absolutely error free, because it is the signalling/control information of the physical link network. So after an error, retransmission is always necessary, this in contrary to standard ATM cells.

8 bits are available for coding only the three states of the idle channel. Thus it has to be possible to decode them correctly when they suffer one or two bit errors. The probability of less then three bit errors can be calculated using the binomium from Newton, thus

\[ P(x \leq 2) = (1 - P_e)^8 + 8(1 - P_e)^7 P_e + 28(1 - P_e)^6 P_e^2. \] (5.13)

The usage of the binomium of Newton implies a negligence of burst errors, i.e. independent error probabilities for all the bits.

Using the property that

\[ (1 - P_e)^n \approx 1 - n P_e \quad \text{for} \quad n P_e \ll 1, \] (5.14)

formula 5.13 can be simplified to

\[ P(x \leq 2) \approx 1 - 28 P_e^2 - 168 P_e^3 = 1. \] (5.15)

Because of this it is assumed that the idle, busy and stop signals can always be detected correctly.

However, errors in the request messages will always make a retransmission necessary. Transmission errors can be detected in two ways.

First, ICMA-CO detects transmission errors in the first 5 slots of the request as a collision. These five slots consist of 40 bits which can be disturbed. The probability for incorrect transmission can again be calculated using the binomium from Newton, which gives

\[ P(X \geq 1) = 1 - P(X = 0) = 1 - (1 - P_e)^{40}. \] (5.16)

Using the property given in formula 5.14., formula 5.16 becomes

\[ P(x \geq 1) \approx 1 - (1 - 40 P_e) = 40 P_e. \] (5.17)

So the probability of an error in these 5 slots can be approximated by 40 \( P_e \).

In this case a stop signal will be transmitted on the idle channel. This means the system moved from the Success State to the Collision State (see figure 5.6).
In figure 5.6 is shown how the new state diagram can be transformed in a similar diagram as figure 5.5. The new state transition probabilities are

\[ P_a = P_r - 40 \cdot P_e, \quad (5.18) \]

\[ P_b = \frac{P_r^2 + 40 \cdot P_e}{1 - P_r} = \frac{P_r^2 + 40 P_r (1 - P_r)}{1 - P_r}, \quad (5.19) \]

and

\[ P_{no} = 1 - P_a - P_b = \frac{1 - 2P_r}{1 - P_r}. \quad (5.20) \]

It is obvious that the \( P_{no} \) is in both models the same. The different \( P_a \) and \( P_b \) indicate a new expression for \( E\{CT\} \),

\[ E\{CT\} = \frac{P_r (Y + 15) - 10P_r^2 + 1 - 10 \cdot 40P_r (1 - P_r)}{1 - P_r}. \quad (5.21) \]

With these expressions the average number of reschedulings \( N_r \) and the throughput \( S \) can be calculated easily (same as in section 5.3.).

Second, the central station can detect an error when the Error Check of the whole message is decided to be incorrect. These errors will be handled by the SRMA protocol: a Time Out is generated and the terminal will reschedule the transmission. Thus, this check detects the transmission errors in the last 11 slots of the request message. The transmission error probability can be calculated in a similar way as above described. However, while in this case 88 bits are transmitted, the approximation will be less accurate. With a \( P_e \) of \( 10^{-6} \) this effect is fortunately very small. So the probability of a transmission error is approximated.
to be $88 \, P_e$. The number of reschedulings $N_{rr}$, caused by these errors, can be calculated similar to formula 5.9. This can be expressed as

$$N_{rr} = \frac{(2-88P_e) \, 88P_e}{1-88P_e} \quad (5.22)$$

With a $P_e = 10^{-6}$, $N_{rr}$ will be $1.76 \, 10^{-4}$.

5.5. Conclusions

In this chapter the performance of the Improved ICMA-CD protocol has been studied. Expressions are derived for both the throughput and the number of retransmissions. Both parameters depend upon the probability of only and only one request attempt in one slot, $P_r$. The packet detection time $Y$ (in slots) is also very important. In the case of a non-ideal transmission channel the Bit Error Rate influences also the performance. $P_r$ depends upon the traffic and the slot time, i.e. the transmission bandwidth. Both the traffic and the bandwidth will be the subject of study in the next chapters.

The search for an appropriate channel allocation mechanism is the aim of this chapter, as introduced in chapter 2. This implies the determination which information channel has to be used by which terminal, and the bandwidth division among the channels and the users. The allocation mechanisms have two main purposes [Van Engelshoven, 1991]. They are used primarily to guarantee the Quality Of Service (QOS) of the connections. Of second importance is the aim to achieve the highest possible multiplexing gain. The statistical gain is the factor by which the sum of the peak bandwidth exceeds the output channel's capacity [Gilbert et al. ,1991].

6.1. The existing channel allocation protocol

The proposal for a radio based ATM protocol will be reviewed and evaluated in this section [Sonnemans, 1990, chapter 6]. This includes the channel allocation mechanism and the queuing model of the whole system.

6.1.1. Protocol description

The transmission of ATM cells through the radio network is the primary task of the protocol. The transmission of single cells is always a part of a session. A session is established using the BTMA multi access protocol on the (SRMA) request channel. The user is given the asked capacity if enough free bandwidth is available. He is never allowed to send more cells during a session than this asked (peak) capacity. In the virtual transmission channel, the reserved slots, are allocated in such a way that the Inter Arrival Time (IAT) between time slots is a constant for a particular user. This is the same procedure as in TDMA transmission systems. However, this isn't mentioned by R. Sonnemans. The central station synchronizes the system by the allocation of the capacity. This is achieved with the broadcast of acknowledgements by the central station. All the terminals receive the acknowledgements. These contain which slots are assigned to a user.

Other active terminals can't use the unfilled slots in a virtual channel of another active user, this capacity is simply thrown away. The time between the arrival of an ATM cell at an active terminal and the actual transmission of that cell, is defined as the cell delay. This delay depends upon the number of previous arrived cells which are still waiting for service and the time between arrival and next allocated slot. The queue of the waiting cells is of the
The distribution of the waiting time (waiting time versus probability) of this queuing system is only known in the s domain.

### 6.1.2. Cell delay

The following analysis is presented in [Sonnemans, 1990] to determine the cell delay $D_c$ (short summary).

1. Time slots are allocated to an user in such a way that the inter arrival time between successive time slots is a constant, $1/\mu_c$.
2. This is a $M/D/I$ queuing system, with $D=1/\mu_c$. Extra delay: time between packet arrival and next allocated time slot $\leq 1/\mu_c$. This gives:
   
   $D (M/D/I) \leq D_c \leq D (M/D/I) + 1/\mu_c.$

   
   s domain: $M/G/I \rightarrow M/D/I$: unable to transform to time domain.
   
   z domain: $M/G/I \rightarrow M/D/I$: unable to transform to time domain. Transformation possible with the intermediate value theorem. This gives a numerical estimation of $P(q=Q)$ ($Q$=queue length).

4. $P(D(M/D/I) \geq (Q+1)/\mu_c) \leq P(q \geq Q)$, $Q$=queue length. In words: the probability the $M/D/I$ delay is longer than $Q+1$ IAT's is smaller than the probability that the queue length exceeds $Q$.
   
   $P(D(M/D/I) + 1/\mu_c \geq (Q+2)/\mu_c) \leq P(q \geq Q)$. The same expression as above.
   
   $P(D_c \geq (Q+2)/\mu_c) \geq P(D_c \geq (Q+2)/\mu_c)$. Use point 2.
   
   $P(D_c \geq (Q+2)/\mu_c) \leq P(q \geq Q)$. Combination of the two last mentioned expressions. This inequality is plotted in [Sonnemans, 1990].

### 6.1.3. Evaluation

The analyzed protocol is a classical TDMA system. However, the already existing knowledge about these kinds of systems (no references concerning this subject) hasn’t been used. The analysis seems to be all right, but it is an approximation.

Strange is that a downlink information channel -from the central station to the remote- is missing. From a communications point of view this is of course a requirement. The handling in the same way of all the traffic streams is another shortage. It is more obvious to treat at least the continuous and bursty traffic in different ways.

The central station synchronizes the system with acknowledgements. How this is done in practice isn’t clear from the report. As last is mentioned that a protocol is missing which
deals with erroneous cell transmissions. This is absolutely necessary because of the (for ATM) high BER of the radio links.

Conclusion is that the protocol of R. Sonnemans is a simple solution for a restricted environment.

In the following parts of this chapter will be tried to develop a more adequate and realistic protocol. To understand the problems, the different facets will be explained in this chapter.

6.2. Network Properties

Some general properties of the indoor IBN are looked at in this section. The description of the required network performance is started with. Next the question will be answered whether statistical or deterministic multiplexing should be used. At last the expected traffic of the network will be the point of attention.

6.2.1. Network performance

Normally users of networks are only interested in the services a network can offer. If someone uses his telephone, then he wants to speak and hear the other party in a 'natural' way. This means that the network should offer a specific Quality Of Service [Gilbert et al,1991]. The ATM layer must accommodate two classes of QOS:

1. With a committed information transfer quality of service per connection. (e.g. telephone, video).
2. With an averaged information transfer quality of service over many connections. (e.g. data services).

In practice this means that a user demands for connections with a specific error rate and a maximum delay.

The network operator has another goal: to provide as many as possible connections with minimal equipment. This means that he asks for a network with a high theoretical throughput. Finding an optimum in the trade off between throughput and delay will be the major point in the design of the network system.

6.2.2. Statistical and deterministic multiplexing.

In ATM the connection resources can be allocated on either the deterministic or statistical multiplexing method. Both types will be explained in this section.

**Deterministic multiplexing** implies that each connection is allocated resources to
Deterministic multiplexing implies that each connection is allocated resources to accommodate its peak cell rate. This means that the sum of the peak bandwidth off all established connections is less (or equal) then the transmission channel capacity. This indicates:
- cell level congestion is totally eliminated.
- negligible cell loss.
- no cell delay variations.
- limited traffic control.
- no bandwidth fragmentation.
- no different hierarchical levels.
Disadvantage is that the transmission capacity is not efficiently used, i.e. there is no multiplexing gain.

Statistical multiplexing implies that the capacity of an output channel is less than the sum of the peak connection bandwidth, but it is larger than their average total bandwidth requirements. Some multiplexing gain is possible. This is most advantageous for high bandwidth variable bit rate services. It uses the law of large numbers, it seems that it is only beneficial when more than 20 terminals use on one link [Gilbert et al, 1991]. The main disadvantages are:
- the finite probability of cell level overload and congestion, which makes buffers necessary.
- variable delays.

It has to be noted that these multiplexing methods are always used in a fully connection oriented ATM environment.

While about 20 variable bit rate users are necessary to reach some statistical multiplexing gain [Gilbert et al, 1991], it is not very suitable to a small local network with a very limited number of users. Because of this, in the IRIBN is chosen for deterministic multiplexing. This is only of importance for the ATM call acceptance mechanism.

6.2.3. Traffic characteristics

To develop a queuing model of the IRIBN it is necessary to model the statistical traffic behaviour of the sources. The traffic generated by the IRIBN is somehow the same as for a broadband LAN. However, although several studies are available which describe some of the statistical characteristics of LAN traffic, a complete description of the statistical behaviour
and the associated traffic models remains to be found. Furthermore, future new services can have a big influence on such a model. The ATM behaviour has to be service independent finally.

The information packets offered to the ATM layer are always generated by higher OSI layer applications. The AAL adapts these to the ATM cell format. On this level four service classes are distinguished, as can be seen in figure 6.1.

![Figure 6.1: Service Classes for AAL](image)

With the three variables, which are the bit rate, connection mode and the timing between source and destination, are theoretically eight combinations possible. The CCITT has only defined the classes corresponding to existing services. These are [de Prycker, 1991]:

**Class A:** A time relation between the source and the destination is needed, the service is connection oriented and the bit rate is constant. Examples are fixed rate video and voice.

**Class B:** The same as class A, except that the bit rates are variable. Examples are variable bit rate video and audio.

**Class C:** A time relation between the source and the destination is not necessary. There are no strict requirements on the network caused delays. An example is connection oriented data transfer and signalling.

**Class D:** This connectionless class is a little bit a stranger in the connection oriented ATM. Connectionless data transport like Switched Multimegabit Data Services (SMDS).

The AAL adapts the packets generated by these sources to the information field of ATM cells, i.e. 48 bytes. The 'adaptation packets' are called Segmentation And Reassembly (SAR) packets. These information 'packets' contain also some fields to guarantee a correct service,
The SAR packets are transmitted transparent through the ATM and Physical Medium layers. Eventually the control information in the SAR packets can be used in the indoor sub network, but it is strictly forbidden to change them.

The precise description of the behaviour of the traffic sources can be subject of further study.

6.3. The basic allocation protocol.

In the previous suggested system was decided to synchronise the TDMA system using separate acknowledgements for each information cell over the acknowledgement channel of the SRMA request protocol [Sonnemans, 1990]. The timing of the scheduling was predetermined for a whole session.

It was already suggested that for the variable bit rate sources (e.g. AAL class B) it would be more efficient when a terminal can ask for more capacity in case of temporally congestion. This was seen as a session (request) problem. In my opinion requests for additive capacity should take place within a session. This because the variance in the bit rate of the source is known before a 'call' is accepted. This indicates a more dynamic/adaptive scheduling of the acknowledgements.

A possible solution is the usage of tokens which are added to the information cells [Malyan, 1991]. Two token types can be distinguished. The first is filled by a remote and is transmitted on the uplink. The contents of the token can be e.g. continue the transmission after this packet or last packet of transmission. The second is filled by the central station and includes the acknowledgements for the terminals (synchronizing the TDMA system). The tokens send from the central station to the remotes should contain also a destination address. This mechanism can be used as a basis for a more complex allocation protocol. It has to be noted that these tokens are used only during a session. This allocation protocol is similar to the Generic Flow Control mechanism for Customers Premises Networks (CPN) in the ATM layer. This protocol is still under study by the CCITT. For the Generic Flow Control are four bits reserved in the cell headers. This means that the cell header (new GFC code word) and the HEC can change during the travelling of a cell through the Customer Premises Network. This indicates more buffer capacity and additive delays. While in chapter 2 is chosen for a sub-network in the physical layer a separate mechanism will still be necessary. However, this is may be an arbitrary decision.
6.4. The Information Channels

The information channels have to support both uplink and downlink channels. Uplink means from the remote to the central station, downlink from the central station to the remote.

In section 2.1. the two major types of connections were described. First the local traffic, second the connections with the wired ATM backbone network. The information channels have to support both kinds of traffic. The further analysis is started with the assumption that the two traffic streams use separated information channels.

Figure 6.2: Local traffic configuration.

The configuration of figure 6.2 can be used for handling the local traffic. In this configuration the downward channel is simply a relayed version of the upward channel, i.e. the information cells/packets which are arriving at the central station are immediately broadcasted. In figure 6.2. TE1 transmits an information packet and an added control token. The central station relays the information and replaces the token. In the next time slot TE3 is permitted to start transmission. The main advantages of a relayed downlink channel are:

- If an uplink is available, all local connections can be made.
- No additive queuing delays in the central station.
- Broadcast and multicast possible.
- Verification by sending terminal possible of own relayed transmission.
- Use downlink channel as synchronisation channel for the remote terminals (use as kind of polling channel).
- Relatively easy implementation.

Disadvantages are:
- Only error control possible on the separate links when the central station buffers the incoming cells/packets. This indicates some additive delays and thus less efficient use of the information channels.

An independent uplink and the downlink is another configuration. For local traffic this will have only disadvantages (always the opposite of the advantages mentioned above).

The possible configurations for the connections with the ATM backbone network are shown in figure 6.3.

Figure 6.3:
Traffic between Indoor Radio Network and backbone ATM Network.

A connection is split up in a radio and a wired part. The wired part uses standard CCITT ATM protocols. This is also the case with the end to end (terminal to terminal) connections.
In figure 6.3. is bidirectional traffic assumed. In (a) is chosen for one uplink and one downlink channel. The uplink channel is now used for the transmission of information cells and tokens. The central station removes the tokens out of the bit stream and gives the ATM cells over to the ATM network. The central station transmits on the downlink channel the reversed traffic and the generated tokens. This means that relaying isn’t possible as described in the previous section.

In (b) is chosen for the same system as in (a) with two more relayed channels offering the same advantages as described above. Which of the two configurations is preferable in this case depends on the used error control mechanism, which will be explained in the next chapter. From the point of consistency, option (b) fixes best with the configuration for the local traffic.

6.5. The Error Control

Additive error checks seem to be necessary on the physical layer because of the relatively bad link quality (and a corresponding high BER) in comparison with optical fibres. In CCITT ATM, the cell header is the only part of the ATM cell which is guaranteed an error free transfer through the ATM layer of the network. This is necessary because the header contains the virtual channel and path identifier which are used for the routing. The headers are protected by a Header Error Code (HEC) which is calculated and added in the physical layer. They are checked on a link by link basis. If an uncorrectable error is detected, the whole ATM cell will be discarded. The AAL verifies the information field of the ATM cells on an end to end basis.

6.5.1. The local error control of the ATM cells

Two scenarios for additive error control are possible in the local radio network. The first solution is to protect the whole ATM cell on a radio link by radio link basis, e.g. with locally introduced error checks. This indicates the use of a data link layer protocol [Tanenbaum, 1988] (OSI layer 2) within the physical layer (see figure 6.4.a). This can be for example a Sliding Window Protocol with window size 1. This is a kind of Stop and Wait Protocol. This is sufficient because of the very short distances and transmission times in the network [Tanenbaum, 1988, pg 225-228].

Doing no additive control is the second solution. This asks for some explanation. One connection will consist in local traffic out of maximum 2 radio links and for outside traffic of only 1 radio link. The radio links are the only links of the connection in the case of the
local traffic. Link by link control will have almost the same results as end to end control in this case. Of course the HEC remains on the ATM layer level (see figure 6.4.(b)).

![Diagram of local traffic with different error control options](image)

(a) Local error control with an added data link layer protocol.

(b) Local traffic without additive error control.

(c) Traffic with backbone ATM network without add. error control.

Figure 6.4: Local Error Control Alternatives.

The radio link will likely be the link with the worst quality for traffic with the ATM backbone network. The BER of the wired ATM will be almost negligible compared to the radio network. Thus end to end error control will not increase the number of retransmissions in the local network. However, this will be probably unacceptable for the operator of the ATM network. This can be avoided when for these kinds of connections the terminal-central station traffic is handled by using the AAL protocols: then end to end error control will be available (see figure 6.4.(c)).

The determination which of the two solutions is preferable will depend on the BER of the radio links. With a high BER it will probably more advantageous to choose for the first option. With a decreasing BER, the second solution tends to become more attractive. From the viewpoint of implementation simplicity the second solution is preferable.

6.5.2. The error protection of the control tokens

It is of course necessary to assure correct transmission of the CPN control tokens besides the error control concerning the ATM cells. Some alternatives will be mentioned.

One option is to check the token together with the ATM cells. This is only possible if is chosen for local error control with a data link layer protocol as described in the previous
Another possibility is to use a protective coding for the tokens and eventually a parity bit. With an appropriate design, the correcting probability of token errors will become almost 100%. So a retransmission scheme isn’t necessary. However, burst errors can destroy this picture. This proposal can be best used in combination with the second option for the local error control. The increase in the system complexity due to the newly introduced error check algorithm is a disadvantage. However, it becomes easy to use, change and access the control token. Less buffering capacity and delay will be the result. This solution looks like a separate token and information channel. In theory two separate channels could be used. However, this has some major disadvantages. First, two channels indicate less efficient use of the bandwidth, e.g. by double timing and transmission delays. Second, two channels need more receiving and sending equipment, making the terminals more expensive.

6.5.3. Choice of total error control.

The choice of an error protecting protocol for the tokens is strongly related with the local error control mechanism. It points out that two combinations can be chosen.

The first option. Use of a data link layer protocol in the ATM Physical Layer on link by link basis. This protocol is used for protection of both the ATM cells and the control tokens. The second option. Use the AAL protocol for protecting the information part of the ATM cells on end to end basis for the entire radio section of a connection. For local traffic this are two radio links, for traffic with the ATM backbone network this is one link. Tokens are protected separately by a protective coding and a parity bit. The second option is chosen.

6.6. The Token System

The main task of the local flow control mechanism is that the accepted calls (by the ATM acceptance system) get their requested QOS. Before introducing the token systems some requirements are made, similar to those for the GFC protocol. These are [CCITT, 1991]:

- Guarantee of capacity. The flow control mechanism must be capable of ensuring that terminals using various bit-rate CBR services, and terminals using VBR services with various traffic sources and having an element of
guaranteed capacity, can be supplied with the capacity assured to them.
- Fair share of capacity. Fair share is defined to mean that any spare capacity is used on a need basis.
- Efficiency. the protocol should not unduly reduce the available capacity.
- The token system should not perform flow control of traffic from the backbone network.
- Because of the use of radio one requirement is added. An error free environment is assumed. This can be done if for example the underlying 'data link layer' protocol operates in a proper way.

There are four types of AAL traffic streams, as can be read in section 6.4. These traffic streams have different service requirements. They get a certain priority to distinguish them, which is used by the flow control system. Four traffic classes at the AAL imply four priorities. However, this isn't necessary. According to [Murase et al, 1991] two priorities can be sufficient.

6.6.1. Token System on the links

The token system takes care of the local transfer of the ATM cells. This has to be done in such a way that the users get their requested QOS. The token system treats the four AAL service classes in different ways. The virtual frames are introduced to assure their quality of service. With the sending of acknowledgements the central station builds the virtual frames. In this way it becomes a sequence of cells from different terminals which operate fully independent of each other. There isn't any timing information within a virtual frame. In general always the same constant length, the virtual frame time $T_f$ s. An example is given in figure 6.5.

![Figure 6.5: Virtual Frames of the token system.](image)
The central station has to schedule the acknowledgements in such a way that the virtual frame structure is used advantageously. A repeatable scheduling and allocation mechanism is most simple and desirable. This means that the allocation mechanism has to be repeated every $T_c$ s. The cell IAT of the slowest class A source is taken as the virtual frame time ($T_c$ s), just as for example POTS in ISDN.

Sources with higher bit rates simply get more 'slots' in this virtual frame than the slower sources. The allocation of more slots in one virtual frame can be done in two ways.

First, the a_tokens can be equally spread over the virtual frame slots. In this way a (near) continuous cell stream is generated. The advantage of such a scheme is that the delay jitter will be minimal. Jitter means that cell displacements in a (continuous) cell stream are small compared to the average cell inter arrival time [Royal PTT, 1991]. This agrees with the ATM standards.

Second, the central station can acknowledge with one a_token a remote station for more than one frame slot. In this way the overhead doesn't increase with the increase of requested capacity. A problem is that the backbone ATM network doesn't accept cell streams with a strong delay jitter as in this case (one could speak even of bursty behaviour). This makes extra buffers necessary between the CPN and the ATM backbone network in the central station. Compared to the first described solution the cell delay doesn't increase in this option [Rom & Sidi, 1991]. It is also a more simple and thus robust protocol, e.g. the timing will become less difficult.

Which of both procedures has to be followed can be subject of further study.

6.6.2. Class A Traffic.

First will be discussed how class A is handled in the token system. This is illustrated by figure 6.6.

After a connection has been established using the request protocol (see chapter 3 and 4), the central station starts sending acknowledgements tokens (a_token) to the remote terminal. After the reception of this a_token, the remote will send its information packet (one or more ATM cells) and an added token, which is normally an continue token (c_token). The central station will then continue with sending acknowledgements to the remote. The time between two a_tokens will be constant like in a TDMA system. When the terminal wants to end the session, it sends an end token (e_token). The central station responds with stopping the scheduling and transmission of a_tokens.

It has to be noted that only if the generated traffic is average a whole number of cells per virtual frame, e.g. 2, the user can receive exactly this capacity per virtual frame. If the cell
rate for instance is 1.5 cells/frame, then altering 1 and 2 cells/frame will be scheduled. During a session this will never change. The central station can predict the needed number of cells/frame. Thus, the transfer of control information isn’t necessary.

6.6.3. Class B traffic.

Class B has almost the same QOS requirements as class A traffic. The difference is that the bit rate is variable. If cells are allocated on the peak cell rate of the source, it is naturally possible to use the same mechanism as for class A. However, here a more sophisticated algorithm will be described.

The starting procedure is the same as for class A. As starting a_token rate the mean bit rate is suggested. During a session two things can happen, an empty or full transmission buffer at the remote terminal, as can be seen in figure 6.7.

When the transmission buffer becomes full, the remote requests for more capacity. This is done by sending a buffer full token (bf_token). The central station can anticipate in two ways.

When the cell rate of the remote decreases again, the transmission buffer will become empty. It will now send a buffer empty token (be_token). The central station will decreases the number of acknowledgements. To avoid too many changes in the scheduling, this last
response is only done after the reception of two successive be_tokens.

6.6.4. Class C&D traffic

The two remaining classes (C and D) differ only in their connection mode. The QOS is for both of them the same. This indicates they need only one priority. They are handled in a totally different way as the above described classes. An example is given in figure 6.8.
The multi access protocol is used for sending a request for the start of a session. When a session is acknowledged, the information transfer can start. The central station schedules the a_tokens in the unused slots of the TDMA frames by the classes A&B, who have always a higher priority. This is continued for as long as the remote has any cells ready for transmission. For the control are again the c_token, e_token, bf_token and the be_token available.

6.7. Control of token mechanism

The central station controls the token system. To do this in a proper way, the central station has to know the state of the system. The subject of this section is how this can be done.

6.7.1. Determination and description of the system state

The allocation mechanism is repeated every virtual frame time \( T_e \) s (see section 6.6.1.). For this purpose the central station uses a table containing the characteristics of all the current
connections. An example is given in figure 6.9.

A traffic table can contain several parameters characterizing the connections. First, the central station has to know the traffic class of a certain connection. This is necessary for knowing the priority of a certain user. Second, it is useful to know the mean traffic (bit rate), the peak traffic (bit rate) and the burstiness [Murase, 1991]. These are also important for the ATM call acceptance mechanism.

Further, it is very important to account how many slots are allocated to each terminal in the virtual frames. The summation of this last column of the traffic table gives the system load in a certain virtual frame. The link load will always be less than or equal to the total link capacity.

Eventually an extra column can be added for still open requests for more capacity. However, this is may be not necessary while congestion is almost excluded (see section 6.6.).

In practice two tables will be necessary, one for the current frame and one for the next frame. This makes the read/write timing much more simple, and is very easy to implement.

<table>
<thead>
<tr>
<th>USER</th>
<th>TYPE (CLASS)</th>
<th>MEAN TRAFFIC</th>
<th>PEAK TRAFFIC</th>
<th>BURSTINESS</th>
<th>ALLOCATED NUMBER CELLS/FRAME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>64 KBIT/S</td>
<td></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>B</td>
<td>9.6 MBIT/S</td>
<td>20 MBIT/S</td>
<td>0.8</td>
<td>170</td>
</tr>
<tr>
<td>5</td>
<td>B</td>
<td>1 MBIT/S</td>
<td>2 MBIT/S</td>
<td>0.4</td>
<td>30</td>
</tr>
<tr>
<td>10</td>
<td>D</td>
<td>-</td>
<td></td>
<td>0.5</td>
<td>50</td>
</tr>
</tbody>
</table>

Figure 6.9: Example of traffic table in the central station.

251 LINK LOAD
(≤ LINK CAPACITY)
It has to be noted that this is not necessary.

The described traffic table in this section offers all the information of the current state of the token system.

6.7.2. Capacity allocation

In this section will be tried to give an answer to the question how the bandwidth of the system should be divided among the active users, i.e. how many cells/slots are allocated to each user in the virtual frames. It has to be emphasized that the ATM call acceptance mechanism guarantees enough capacity. This means that under normal circumstances the token system is never moved in a congestion state. In case of emergencies, the token system has to be stable. Stable means an autonomous return to normal operation after e.g. a temporary congestion.

Class A traffic is relative easy to handle. If an user has been accepted, and is in an information transfer session, it needs each virtual frame the same amount of capacity. During a session this will never change. When the cell per frame rate is not a whole number, a more complex situation occurs (see section 6.6.2). All the class A traffic is scheduled in the first part of the virtual frame, see figure 6.5.

Because the class B traffic has a variable bit rate and thus a cell/frame rate, it is more difficult to handle. A session starts with the allocation of the number of cells/frame corresponding with the average bit rate of the VBR source.

Imagine that an user has a temporarily need for more capacity. It transmits a request for more capacity (bf_token). The central station calculates the available capacity for this class B user, i.e. the link capacity minus the capacity allocated to the other class B and A users in the current frame. If the requested capacity fits in the available bandwidth, the user is serviced. If not, cells have to be queued or discarded at the VBR source.

In practice, the sum of the requested extra capacity of all Class B sources (instead of the capacity of the single user will be used). The available bandwidth is fairly divided among the requesting users, as demanded in section 6.6. This is in theory preferable.

A major problem is that the system includes only the transmission buffer full token. But how much extra capacity is needed? There are a few options.

First option is to introduce more sophisticated tokens. These tokens could represent the requested capacity. But how can the remote terminal predict the future? It can make only an approximation on basis of its traffic characteristics and parameters. But these are also known
by the central station (traffic table).

This brings us to the second option. The central station tries to estimate the needed extra capacity in case of a received bf_token. This is done with the collected parameters in the traffic table. From these parameters a capacity multiplication factor $C_{mpf}$ can be calculated. This can have as basis the quotient of the peak bit rate $b_p$ and the currently allocated bit rate $b_{ca}$, thus

$$C_{basic} = \frac{b_p}{b_{ca}} .$$

(6.1)

If this quotient is taken as capacity multiplication factor, the source is always allocated its peak bit rate in the next frame. However, this can lead to a waist of capacity. Because of this the burstiness is taken into account. The normalized burstiness factor $BF_n$ is introduced for this purpose. This factor has a value between 0 and 1, indicating not bursty and very bursty. Suggestions for the determination of this factor can be find in [Royal PTT, 1991]. With this burstiness factor taken into account the capacity multiplication factor becomes

$$C_{mpf} = (C_{basic} - 1) \cdot BF_n + 1 .$$

(6.2)

With this formula the allocated bit rate in the next virtual frame becomes

$$b_{ca, next} = b_{ca, current} \cdot C_{mpf} .$$

(6.3)

Take for example a traffic source with a peak bit rate of 10 Mbit/s and a mean bit rate of 5 Mbit/s. Currently it is 8 Mbit/s allocated. This means that $C_{basic}$ becomes $10/8 = 1.25$. If it is a bursty source, e.g. $BF_n = 0.9$, the $C_{mpf}$ becomes 1.225. The $b_{ca}$ will thus become in the next slot $8 \times 1.225 = 9.8$ Mbit/s. A less bursty user with a $BF_n = 0.2$ will get a $C_{mpf} = 1.05$ and the $b_{ca,next} = 8.40$ Mbit/s.

In case of a request for a decrease of the allocated capacity (be_token) another scheme has to be used. The $b_{ca,next}$ has now to be a part of the previous allocated bit rate. It is very simple to use for the $C_{mpf}$ the 'complement' of the burstiness, i.e.

$$C_{mpf} = 1 - BF_n .$$

(6.4)

This formula is used together with 6.3. in a same way as described in the above example. Class C and D traffic doesn't almost has delay restrictions, it has always a low priority. They are acknowledged only if capacity is available after all the class A and B traffic has been serviced. All this capacity can be used by the class C & D traffic. From the viewpoint of fairness, all the sources share the available capacity equally. This is assured by the central
station. In case of heavy network load, this can lead to huge delays for the users in these classes.

The described procedure is a simple solution for the capacity allocation problem on the transmission links. It is flexible and can deal with different traffic sources. The response time of the system is at least the virtual frame time $T_c$. As last problem remains how the system should deal with the signalling information. This will be discussed in the next section.

6.7.3. Signalling, which priority?

In ATM the signalling is performed using separated ATM channels. Two stages can be distinguished in the signalling process. The signalling starts always with the use of the Meta Signalling Channel (MSC). This channel has a predetermined VCI and VPI which is the same for the whole ATM network, and is used by all the network users. Using the MSC, the terminal is assigned to a specific Signalling Virtual Channel Link (SVCL) by the ATM network. A SVCL belongs always to one ATM connection. This channel is used for e.g. negotiating about the wanted QOS. The signalling channel is of the Class C type [de Prycker, 1991]. A maximum permitted delay for the SVCL channels is about 260 ms (up and down link, value is taken from the ISDN). The question is now, how to deal with the signalling cells.

Signalling is considered as important in a network. It should be avoided that, in case of a very heavy load, signalling cells will be lost or delays become excessive.

There are some alternatives.

First, the signalling traffic is treated the same as normal Class C traffic. This is the lowest traffic priority. This means that in a congestion situation the delays can become too high. However, this will be very rare with a good working ATM call acceptance mechanism.

Another possibility is to introduce a kind of C+ class. In this class cell allocation is assured even in heavy load circumstances (of course only in a session!!). This is always done by giving these connections at least every 130 ms the highest priority. The central station manages this using the class type entry in the traffic table.

This C+ class can also be treated in a total different way then the normal C class traffic. Signalling messages are normally very short, maximums are mentioned of about 6 cells (300 bytes in ISDN). If a signalling session is started, the token system gives within 130 ms an acknowledge for the whole message. This is always done in one virtual frame, after which the session is immediately terminated. A terminal is acknowledged for the maximum signalling message length of 6 cells. This reduces the system complexity and assures in all
situations the transfer of the signalling system.

It seems that the two last mentioned solutions are preferable. The solution which treats the signalling traffic as ordinary C traffic with, in case of heavy system loads, an assurance for cell transfer will use the available bandwidth efficient. System complexity will increase due to this mechanism. The solution in which a whole signalling messages is served in one virtual frame is less complex. However, it will waist some system capacity. The conclusion is that both solutions are acceptable. Which of the two is chosen will be a trade off between system complexity and bandwidth efficiency.

6.8. Queuing analysis

In this section will be looked at the M/D/1 queuing system and the TDMA system. As described in section 6.1., the cell delays in the IRIBN can be determined with these systems. In this analysis is assumed that the inter arrival time between two allocated slots is constant. This is typical for class A traffic.

6.8.1. M/D/1 queuing system

The M/D/1 queuing theory is a special case of the M/G/1 queuing systems. Aim is to find the expressions for the waiting times in the system. Both average and (cumulative) distribution formulas for the waiting system are looked at. Most of this theory is from [Kleinrock, 1975] and [Gross & Harris, 1985]. It has to be noted that all the expressions are only valid for the time/point just after a customer left the queue. In this section the transmission packets will be denoted by customers, which is usual in the queuing theory.

The analysis starts always with the equation of Pollaczek-Khintchin transform equation, which gives the z transform of the number of customers in the system, \( P(q=Q) \), (\( \rho \) is the system utilization, [Kleinrock, 1975,pg.194]),

\[
Q(z) = \frac{(1-\rho)(1-z)}{1-ze^{\rho(1-z)}}.
\]  

(6.5)
From this z transform the following expressions for $P(q=Q)$ can be derived [Gross & Harris, 1985, pg. 269],

$$
P(q=0) = 1 - \rho \\
P(q=1) = (1 - \rho)(e^\rho - 1) \\
P(q=n) = (1 - \rho) \sum_{k=1}^{n} (-1)^{n-k} e^{k\rho} \left[ \frac{(k\rho)^{n-k}}{(n-k)!} + \frac{(k\rho)^{n-k-1}}{(n-k-1)!} \right] \quad (n \geq 2)
$$

(6.6)

These last expressions were evaluated with a computer and plotted in figure 6.10.

Figure 6.10: $P(q=Q)$ for M/D/1 system for $Q=0$ to $9$.

With formula 6.5 it is possible to calculate the distributed cumulative expressions for the queue lengths. For $P(q \geq Q)$ is find

$$
P(q \geq Q) = 1 - \sum_{1}^{Q-1} P(q = Q) .
$$

(6.7)

This expression is plotted in figure 6.11.
For the cumulative distribution $P(q < Q)$ has been found

$$P(q < Q) = \sum_{i}^{Q-1} P(q = Q).$$  \hfill (6.8)

This expression is plotted in figure 6.12.

Figure 6.12: $P(q < Q)$ for M/D/1 system with $p$ and $Q$ as parameter.
Besides these expressions, concerning the number of customers in the system on one service moment, some expressions are known for the averages of some system variables.

First, the expression for the average queue length $\bar{q}$ is

$$\bar{q} = \frac{2\rho - \rho^2}{2(1-\rho)} , \quad (6.9)$$

which is plotted in figure 6.13 [Kleinrock, 1975, pg. 188].

![Figure 6.13: Average M/D/1 queue length.](image)

Compare this result with the result of figure 6.10. The maxima in the curves of 6.10 ($P(q=Q)$) seem to correspond with the points of the curves in figure 6.13. The average number of customers in the waiting queue increases very fast for $\rho$ smaller than 0.2 and higher than 0.8.

The average time spent in the system, $S$, normalized to the service time ($\bar{x}$, corresponding to $T_e$), is

$$\frac{S}{\bar{x}} = \frac{1 - \rho}{2} \frac{1-\rho}{1-\rho} . \quad (6.10)$$

which is plotted in figure 6.14.
From formula 6.10 it is easy to derive the expression for the average waiting time, which is the total time spent in system minus the service time,

\[
\frac{W}{\bar{x}} = \frac{S}{\bar{x}} - 1 = \frac{\rho}{2(1-\rho)} .
\] (6.11)

The last part of this section deals with the distribution of the waiting time in the system. This is done using the Laplace transform. The Laplace transform of the distribution of the waiting time is denoted by \( W^*(s) \), the transform of the distribution of the total time spent in the system is \( S^*(s) \). In [Kleinrock, 1975, pg. 199-200] is for the M/D/1 queuing system derived that

\[
S^*(s) = e^{-sr} \cdot \frac{s(1-\rho)}{s - \lambda + \lambda e^{-sr}} .
\] (6.12)

and

\[
W^*(s) = \frac{s(1-\rho)}{s - \lambda + \lambda e^{-sr}} .
\] (6.13)

It is complicated to transform these expressions back to the time domain [Sonnemans, 1990]. We will use these expressions in the next section to compare the M/D/1 system with TDMA.
6.8.2. TDMA queuing model.

In [Rom & Sidi, 1990] is presented a queuing analysis of the TDMA queuing system. A lot of results are presented, but most of them are very difficult to understand. In this section we will briefly discuss some results.

The delay suffered by packets in TDMA systems has three components: (1) the time between its generation and the end of the current frame, (2) the queuing time to allow all the packets already queued to be transmitted and (3) the packet transmission itself. In the TDMA system the duration of a frame is always $T_c$ s, the packet transmission takes $T$ s, and there are $M$ slots of $T$ s in one frame (thus $M$ users!), as can be seen in figure 5.15.

![Figure 6.15: TDMA queuing system.](image)

The total expected packet delay is

$$D = \frac{1}{2} T_c + W_q + T = \frac{\rho MT}{2(1-\rho)} + T + \frac{MT}{2}$$

$$= T \left[ 1 + \frac{M}{2(1-\rho)} \right]. \quad (6.14)$$

It is again difficult to derive expressions for the message delay distribution. In this section will be looked at a model with only messages of 1 packet. For the special Laplace transform (see previous section) of the packet delay distribution is found that

$$D^*(s) = \frac{1-\rho}{T_c} \frac{1 - e^{-sT_c}}{s - \lambda + \lambda e^{-sT_c}} e^{-sT_c} e^{-s(\sigma_s - \eta)} e^{-sT_c}. \quad (6.15)$$

If for the packet transmission time $T$ is taken the same time as the frame time $T_c$, a system
has been created which is comparable to the M/D/1 system. But how comparable? Are both systems may be the same?

Substituting $T_c$ for $T$ in formula 6.15 gives

$$D^*(s) = \frac{(1-\rho) e^{-sT_c}}{s-\lambda + \lambda e^{-sT_c}} \frac{1-e^{-sT_c}}{T_c}$$ \hspace{1cm} (6.16)$$

If both systems would be the same $D'(s)$ and $S'(s)$ (6.12 and 6.16) have to be equal. Comparing both formulas this doesn’t seem to be so. This is only the case when

$$\frac{1-e^{-sT_c}}{T_c} = s.$$ \hspace{1cm} (6.17)

To solve 6.17 we use the McLaurin series for $e^x$, which gives

$$e^{-sT_c} = \sum_{n=0}^{\infty} \frac{(-sT_c)^n}{n!} = 1 - sT_c + (sT_c)^2 - (sT_c)^3 - ....$$ \hspace{1cm} (6.18)

If the second and higher order terms are neglected it is immediately clear that the equation 6.17 becomes true. Remains as question when these higher order terms may be neglected. While the equality has to be true for all $s$, it is necessary that $T_c$ is very small. This means that the TDMA system and the M/D/1 queuing system behave the same when the frame time is small. With frame times of the order of $\mu s$ this will be the case.

But why are both systems are under normal circumstances different? This is illustrated with figure 6.16.

![Figure 6.16: M/D/1 and TDMA example.](image)

CELL ARRIVALS AND EXPECTED NUMBER OF PASSED FRAMES BEFORE SERVICE

Figure 6.16: M/D/1 and TDMA example.
In the first $T_c$ s five new packets arrive. For the M/D/1 system these are seen at the beginning of the new frame, the queue size increases then with five. In the calculation by [Rom & Sidi, 1990] the delays are calculated per arrival. This means that the position in the sequence of arrivals within a frame affects the delay distribution. The time from arrival and begin of the next frame is taken into account. This gives per arrival an added delay between $0$ and $T_c$. This explains the above derived result that both distributions are almost the same for small $T_c$ values.

The conclusion of this section is that the TDMA analysis presented in [Rom & Sidi, 1990] offers in a environment with very short frame times the same results as for the M/D/1 queuing system. Thus it can be sufficient to know the characteristics of the M/D/1 queuing system to predict the delays in the network.

6.9. Conclusions.

In this chapter was searched for a protocol for the information transfer and the channel allocation algorithm. The suggested protocol by Sonnemans [1990] is a solution for a restricted environment. The radio network has to assure the Quality of Service (QOS) of the connections. Deterministic multiplexing is preferable above statistical multiplexing in the indoor IBN. Four AAL traffic classes have to be transferred. A token system is applied as bandwidth allocation protocol. The AAL performs the additive error control. A separate check protects the tokens. The central station controls the token system.

An algorithm for capacity allocation has been suggested. This is a first order control system. The signalling traffic is treated as a separate traffic Class. The TDMA system problems can be approximated by M/D/1 queuing systems in this system. This can be used to calculate the cell delay of class A traffic.
7. System Proposal

The proposed IRIBN is an ATM compatible network. The adaptation of the radio network to standard ATM is performed in the transmission frame adaptation function of the physical layer. The IRIBN is an extended star network with the concentrators in one physical entity with the central station. The available bandwidth for the system is divided into control and information channels. The control channels ensure efficient use of the information channels.

Two levels are distinguished within an ATM call. The connections are established in the ATM and higher layers. The ATM call acceptance mechanism decides which terminal is given a connection. The traffic table in the central station is filled in this ATM call level. When a terminal wants to start cell transmission, it moves to the session level. A session is defined as a burst/sequence of ATM cells transmitted within an ATM call. Session request are made using the improved idle casting multi access collision detect protocol (I-ICMA-CD). This protocol is used in the control channels. A token system takes care of the information transfer and the channel allocation within a session. This system is used within the information channels. Each of the AAL traffic classes and the signalling have their own priority. In the token system 'tokens' are piggy backed on the information packets. These tokens contain the control information of the system. There are tokens for continuing of transmission, end of transmission, buffer full, buffer empty and acknowledgements. A preventive error coding and a parity bit protect the tokens against bit errors. The information part of the messages isn't protected in the token system. The ATM Header Error Checks and the ATM adaptation layer protect them against errors. The central station controls the token system. The token system has a repeatable allocation mechanism. This is done per virtual frame. The number of allocated slots in a virtual frame to a particular terminal depends on the system and terminal state. The central station knows the system state from the traffic table. The terminal state is transmitted by the terminals using the tokens (e.g. buffer full token). These states are used in a first order control system for the capacity allocation.

Two delays can be distinguished in the IRIBN. The session delay is the time between the request for the start of a session and the first acknowledgement for a cell transmission. The cell delay is the time a cell arriving at a terminal has to wait for service.

68
8. Conclusions

The estimation of the chances for realising an IRIBN in the 60 GHz band is the subject of this feasibility study. Emphasis should lay on the fact that the radio links are still studied. Their qualities are still uncertain. Especially in the area of the possible bit rates, the packet detection delay and the Bit Error Rate.

The realisation of the IRIBN looks to be possible. A (technical) proposal has been done for such a network. It is an ATM compatible network. It can be used also in a different context.

The economical aspects have only been briefly discussed. The use of ATM implies that the network will be compatible with the future world standard for broadband networks. The costs of the system have not yet been calculated.

During the study it became clear that the currently developed indoor radio networks, like Digital European Cordless Telecommunications (DECT), are still pre ISDN networks. The exchange of the IBN for an ISDN can be an interesting option. The proposed radio system can still be used then, but the requirements will be less.

A lot of further study remains to be done. Inter room networks have not been studied. This is of particular interest for commercial networks. Further, it hasn't been possible to make an exact queuing model or simulation of the network because of the lack of knowledge of the radio links.
References

[ Abramson, 1985 ] Norman Abramson
Development of the ALOHANET

Performance analysis of ICMA/CD multiple-access for a packet mobile radio.

Analytical Model of Busy Channel Multiple Access (BCMA) for Packet Radio Networks in a Local Environment.

[ CCITT, 1991 ] CCITT
Report on ATM aspects, Annex 2
Report of working party 18/8,
COM 18-R 70-E
June 1991.

[ Choi, 1989 ] Choi, E., Kim, M.
On improving throughput in packet mobile radio network.

[ van Engelshoven, 1991 ] Rob van Engelshoven
Multiserv: an ATM based residential SPN for the Belgian broadband project.

70
Developing a cohesive traffic management strategy for ATM networks.

[Gross & Harris, 1985] Donald Gross and Carl M. Harris
Fundamentals of queuing theory

[Kleinrock, 1975] Leonard Kleinrock
Queuing Systems. Volume 1: Theory

[Kleinrock, 1985] Leonard Kleinrock
On queueing problems in random access communications

[Lee, 1989] Lee, D., Choi, E., Kim, M.
Performance analysis of virtual time ICMA/CD.

[van der Linden, 1992] A. van der Linden
Concepts for shared medium access in ATM networks.
Graduation report, Eindhoven University of Technology.
June 1992, to be published.

A speech packet protocol for mobile radio using idle signal casting multiple access.
Virtual time CSMA: why two clocks are better than one.

[Murase, 1991] Tutomo Murase, Hiroshi Suzuki, Takao Takeuchi
A call admission control for ATM networks based on individual
multiplexed traffic characteristics

Idle-Signal Casting Multiple Access with Collision Detection
(ICMA-CD) for Land Mobile Radio.
IEEE Trans. on Vehicular Technology, Vol.36, No.2 (May 1987),
pg. 45-50.

Idle-Signal Casting Multiple Access with Data Slot Reservation
(ICMA-DR) for Packet Radio Communications.
IEEE Trans. on Vehicular Technology, Vol.38, No.2 (May 1989),
Pg. 50-54.

Control Channel Traffic Design in a High-Capacity Land Mobile
Telephone System.
IEEE Trans. on Vehicular Technology, Vol.27, No.4 (November

[Pach, 1989] Pach, Andrzej R.
Performance study of SRMA multiple access protocols.
Int. Journal of Systems Science, Vol.20 (1989), No.6, Pg. 1053-
1065.
[de Prycker, 1990] Martin de Prycker
Asynchronous Transfer Mode, solution for broadband ISDN.

[Rom & Sidi, 1990] Raphael Rom and Moshe Sidi
Multiple Access Protocols, performance and analysis.

[Royal PTT, 1991] Royal PTT Nederland NV
Characterization of cell level 'burstiness' in ATM.

[Song, 1989] Song, J., Kim, M.
Improved ICMA/CD protocol in packet mobile radio networks.

[Sonnemans, 1990] Rob Sonnemans
ATM in an indoor radio network: a feasibility study.
TUE afstudeerverslag, fac. der elektrotechniek,
Eindhoven, augustus 1990.

[Tanenbaum, 1988] Andrew S. Tanenbaum
Computer Networks
1988, Englewood Cliffs, New Jersey, USA.

Packet Switching in Radio Channels: Part 2- The Hidden Terminal
Problem in Carrier Sense Multiple-Access and the Busy-Tone
Solution.
1417-1433.
Packet Switching in Radio Channels: Part 3- Polling and (Dynamic) Split-Channel Reservation Multiple Access.

Multiaccess protocols in packet communications systems.
List of abbreviations

AAL  ATM Adaptation Layer.
ATM  Asynchronous Transfer Mode.
BCMA Busy Channel Multi Access.
BER  Bit Error Rate.
BTMA Busy Tone Multi Access.
CBR  Continuous Bit Rate.
CCITT International Consultative Committee for Telecommunications and Telegraphy.
CD   Collision Detect.
CPN  Customers Premises Network.
CS   Collision State.
CSMA Carrier Sense Multi Access.
CSMA-CD CSMA Collision Detect.
CT   Cycle Time.
DECT Digital European Cordless Telephone Telecommunications.
DPR  Data Packet Request.
ES   Empty State.
GFC  Generic Flow Control.
FDMA Frequency Division Multi Access.
FIFO First In First Out.
HDTV High Definition Tele Vision.
HEC  Header Error Coding.
IAT  Inter Arrival Time.
IBN  Integrated Broadband Network.
ICMA Idle Casting Multi Access.
ICMA-CD ICMA Collision Detect.
I-ICMA-CD Improved ICMA Collision Detect.
I-IBN Indoor Integrated Broadband Network.
IRIBN Indoor Radio Integrated Broadband Network.
ISDN Integrated Services Digital Network.
LAN  Local Area Network.
M/D/1 Exponential distribution arrivals, Deterministic distribution service, 1 server.
M/G/1  Exponential distribution arrivals, General distribution service, 1 server.
MSC  Meta Signalling Channel.
N-ISDN  Narrowband Integrated Services Digital Network.
NTT  Nippon Telegraph and Telephone company.
OSI  Open Systems Interconnection.
PABX  Private Automatic Branch eXchange.
PM  Physical Medium.
POTS  Plain Old Telephone Services.
QOS  Quality Of Service.
RAM  Request Answer Message.
RM  Request Message.
SAR  Segmentation And Reassembly.
SDH  Synchrone Digital Hierarchy.
SMDS  Switched Multimegabit Data Services.
SRMA  Split Channel Multi Access.
SS  Success State.
STM  Synchronous Transfer Mode.
SVCL  Signalling Virtual Channel Link.
TC  Transmission Convergence.
TDMA  Time Division Multi Access.
TE  Terminal.
VBR  Variable Bit Rate.
VCI  Virtual Channel Identifier.
VPI  Virtual Path Identifier.
VT  Virtual Time.
List of Variables.

\[ \eta \quad b_a = b_r. \]
\[ \eta_a \quad b_a / b_m. \]
\[ \eta_r \quad b_r / b_m. \]
\[ \theta \quad \text{fraction of } W \text{ assigned to the message channel.} \]
\[ \lambda \quad \text{average arrival rate.} \]
\[ \rho \quad \text{system utilization factor.} \]
\[ \tau \quad \text{propagation delay time.} \]
\[ A \quad \text{slot size.} \]
\[ a \quad \text{ratio of the propagation time, between central station and terminal, and the transmission time of a message packet.} \]
\[ B \quad \text{preamble in slots.} \]
\[ BF_n \quad \text{normalized burstiness factor.} \]
\[ b_a \quad \text{length of the answer packets.} \]
\[ b_c \quad \text{'busy' on busy channel.} \]
\[ b_{ca} \quad \text{allocated bitrate.} \]
\[ b_m \quad \text{length of the message packets.} \]
\[ b_p \quad \text{peak bitrate.} \]
\[ b_r \quad \text{length of the request packets.} \]
\[ C_{mof} \quad \text{capacity multiplication factor.} \]
\[ c \quad \text{speed of light.} \]
\[ ci \quad \text{correct investigation of the busy channel.} \]
\[ D \quad \text{TDMA (expected) packet delay.} \]
\[ D^*(s) \quad \text{laplace transform of distribution of TDMA packet delay.} \]
\[ D_1 \quad \text{delay 1.} \]
\[ D_2 \quad \text{delay 2.} \]
\[ D_p \quad \text{packet detection delay time.} \]
\[ D_{pq} \quad \text{packet queuing time.} \]
\[ D_{rq} \quad \text{request queuing time.} \]
\[ D_x \quad \text{random retransmission delay.} \]
\[ E\{CT\} \quad \text{expectation of cycle time.} \]
\[ fc \quad \text{'free' on busy channel.} \]
ratio of the size of the message packets and the polling packets.

number of terminals.

number of slots in TDMA frame (chapter 5) (equals number of terminals).

m th terminal.

number of retransmissions.

average number of reschedulings.

number of reschedulings with errors.

probability of one request in the ES.

probability of two or more in ES.

collision probability.

probability of correct transmission.

characteristics of the transmission channel.

Bit Error Rate of the channel.

inside permission to send probability.

probability there will occur no collisions.

probability there are no request attempts.

probability there is no successful transmission.

outside permission to send probability.

probability a terminal is permitted to send.

probability of request in a slot.

probability there are no reschedulings.

probability of successful transmission through the channel.

z-transform of number of customers in system.

average queue length.

throughput.

time spent in (M/D/1) system.

laplace transform of distribution of time spent in M/D/1 system.

sensing time of busy channel.

packet Transmission Time.

virtual Frame Time.

time needed to transmit a request.

Time-Out.

total available bandwidth.

(average) waiting time in M/D/1 system.
\( W(s) \) laplace transform of waiting time distribution in M/D/1 system.

\( W_a \) bandwidth of the answer channel.

\( w_i \) wrong investigation of the busy channel.

\( W_m \) bandwidth of the message (information) channel.

\( W_q \) waiting time in queue for TDMA.

\( W_r \) bandwidth.

\( W_r \) bandwidth of the request channel.

\( Y \) preamble time.