An ATM adaptation layer for signalling

van den Elsen, R.A.

Award date:
1991

EINDHOVEN UNIVERSITY OF TECHNOLOGY
FACULTY OF ELECTRICAL ENGINEERING
TELECOMMUNICATION DIVISION

AN ATM ADAPTATION LAYER FOR SIGNALLING

by R.A. van den Elsen

Report of graduation work
performed December 1990 to August 1991

Supervisor: Prof. ir. J. de Stigter

Acknowledgements

During my graduation work I had several meetings with members of the 'Communicatietechniek' group of PTT Research.

I want to thank C. Gabriël, P. Blankers and K. Kerkhof (PTT Research) for their assistance and pleasant cooperation during these meetings.

Summary

This report describes my graduation work performed at the Faculty of Electrical Engineering of the Eindhoven University of Technology.

In the Broadband-Integrated Services Digital Network (B-ISDN) is Asynchronous Transfer Mode (ATM) the transfer mode. On top of the ATM layer an ATM Adaptation Layer (AAL) is defined for adapting the characteristics of ATM to various applications. One of these applications is signalling. The goal of my graduation work was to develop such an AAL for signalling.

In the ISDN there are two different signalling protocols (access signalling and network signalling). It is important that an AAL for signalling is common for both the UNI and NNI. This approach is consistent with long term signalling requirements for unified UNI/NNI signalling. The proposed AAL for signalling therefore replaces the functionality of both LAPD and MTP level 2.

Cell loss rate of the ATM layer and random bit errors at the Physical Layer are important parameters that influence the AAL. Causes of cell loss are random bit errors, burst errors and congestion.

The functions of the AAL are separated into functions for the Convergence Sublayer (CS) and functions for the Segmentation and Reassembly sublayer (SAR). The coding of the CS-PDU is based on the HDLC protocol. An uniform state transition diagram based on LAPD is defined for both the UNI and NNI.

The performance of the error correction function is enhanced by the use of selective retransmission. The performance of the AAL (delay, loss, undetected errors) will be much better than the performance objectives of the MTP.

During the standardization of future higher layer signalling protocols effort have to be made to ensure that most frequently used signalling messages can be transported by one ATM cell.

Samenvatting

Dit verslag is geschreven in het kader van mijn afstudeeronderzoek aan de Faculteit der Elektrotechniek van de Technische Universiteit Eindhoven.

In het Broadband Integrated Services Digital Network (B-ISDN) is Asynchronous Transfer Mode (ATM) de transfer mode. Boven de ATM laag is een ATM Adaptation Layer (AAL) gedefinieerd om de karakteristieken van de ATM laag aan te passen aan verschillende applicaties. Eén van die applicaties is signalering. Het doel van mijn opdracht was het ontwerpen van een AAL voor signalering.

In het ISDN zijn er twee verschillende signaleringsprotocollen (UNI en NNI). Het is belangrijk dat een AAL voor signalering uniform is op de UNI en de NNI. Deze benadering is consequent gelet op de signaleringseisen in de toekomst, waarbij gestreefd wordt naar uniforme signalering op de UNI en de NNI. De voorgestelde AAL zal daarom de functie van zowel LAPD als MTP level 2 vervangen.

Cel verlies op de ATM laag en bit fouten op de fysische laag zijn belangrijke factoren die van invloed zijn op de AAL. De oorzaken van cel verlies zijn random bit fouten, burst fouten en congestie.

De functies van de AAL zijn opgesplitst in functies van de Convergence Sublayer (CS) en functies van de Segmentation and Reassembly (SAR) sublaag. De codering van de CS-PDU is gebaseerd op het HDLC protocol.

De prestatie van de functie 'fout correctie' is verbeterd door het gebruik van selectieve hertransmissie. De prestatie van de AAL (vertraging, verlies, niet opgemerkte fouten) zal veel beter zijn dan de prestatie van het MTP.

Tijdens de standaardisatie van toekomstige signaleringsprotocollen voor de hogere lagen moet ervoor worden gezorgd dat de signaleringsberichten die het meest worden gebruikt door één ATM cel kunnen worden getransporteerd.

Contents

List of Abbreviations and Symbols

AAL	ATM Adaptation Layer
ARQ	Automatic Repeat Request
ATM	Asynchronous Transfer Mode
B-ISDN	Broadband Integrated Services Digital Network
BER	Bit Error Rate
BIB	Backward Indicator Bit
BOM	Begin Of Message
BSN	Backward Sequence Number
C/R	Command/Response
CATV	Cable Television
CBR	Constant Bit Rate
CCITT	International Telegraph and Telephone Consultative Committee
CIR	Cell Insertion Rate
CK	Check Bits
CLP	Cell Loss Priority
CLR	Cell Loss Rate
COM	Continue Of Message
CRC	Cyclic Redundancy Check
CS	Convergence Sublayer
CSSN	Convergence Sublayer Sequence Number
DISC	Disconnect
DL	Data Link
DLCI	Data Link Connection Identifier
DM	Disconnected Mode
DSS	Digital Subscriber Signalling
DUP	Data User Part
EM	Error Monitoring
EOM	End Of Message
ES	Errored Second
F	Flag
FCS	Frame Check Sequence
FEC	Forward Error Correction
FIB	Forward Indicator Bit
FISU	Fill In Signal Unit
FRMR	Frame Reject
FSN	Forward Sequence Number
GFC	Generic Flow Control
HDLC	High Level Data Link Control
HEC	Header Error Control
I	Information
ISCP	ISDN Signalling Control Part

ISDN	Integrated Services Digital Network
LAPD	Link Access Protocol on the D-channel
LI	Length Indication
LSSU	Link Status Signal Unit
M	Modifier function bit
MF	Multi Frequency
ms	millisecond
MSU	Message Signal Unit
MTP	Message Transfer Part
N-ISDN	Narrowband Integrated Services Digital Network
NNI	Network Network Interface
NP	Network Performance
NSDU	Network Service Data Unit
NSP	Network Service Part
OAM	Operation and Maintenance
OSI	Open Systems Interconnection
P/F	Poll/Final
PCR	Preventive Cyclic Retransmission
PDU	Protocol Data Unit
PL	Physical Layer
PM	Physical Medium
PRM	Protocol Reference Model
PT	Payload Type
QOS	Quality Of Service
REJ	Reject
RES	Reserved
RNR	Receive Not Ready
RR	Receive Ready
S	Supervisory
S	Supervisory function bit
SABMM	Set Asyncronous Balanced Mode Mod256
SAP	Service Access Point
SAPI	Service Access Point Identifier
SAR	Segmentation and Reassembly
SCCP	Signalling Connection Control Part
SDU	Service Data Unit
SF	Single Frequency
SF	Status Field
SIF	Signalling Information Field
SIO	Service Information Octet
SLS	Signalling Link Selection Code
SN	Sequence Number
SREJ	Selective Reject
SS	Signalling System

SSM	Single Segment Message
ST	Segment Type
STD	State Transition Diagram
STP	Signalling Transfer Point
TC	Transmission Convergence
TC	Transaction Capabilities
TE	Terminal Equipment
TEI	Terminal Equipment Identifier
TUP	Telephone User Part
U	Unnumbered
UA	Unnumbered Acknowledgement
UI	Unnumbered Information
UNI	User Network Interface
UP	User Part
VC	Virtual Channel
VCC	Virtual Channel Connection
VCI	Virtual Channel Identifier
VP	Virtual Path
VPC	Virtual Path Connection
VPI	Virtual Path Identifier
XID	Exchange Identification

τ	One way propagation delay
a	traffic loading
E_p	percentage of errored signal units
M	Cycle of N(R) numbering
N(R)	Receive Sequence Number
N(S)	Send Sequence Number
N_e	maximum consecutive errored signal units
R	transmitted CS-PDUs per second
T_c	cell assembly/disassembly delay
T_{cs}	STP function delay
T_d	cell buffering delay
T_e	delay due to CODEC
T_{mr}	MTP receiving function delay
T_{ms}	MTP sending function delay
T_n	transfer delay (waiting times)
T_p	propagation time
T_q	Queuing delay
W	window size

List of Figures and Tables

Introduction

The emergence of high-speed telecommunications technologies has provided an impetus for the definition of Broadband ISDN. B-ISDN is conceived as an all-purpose digital network. It will provide an integrated access that will support a wide variety of applications for its customers. B-ISDN supports on demand, semi-permanent and permanent, point-to-point and point-to-multipoint connections and provides on demand, reserved and permanent services.

The transfer mode chosen by CCITT Study Group XVIII as the basis of B-ISDN is called Asynchronous Transfer Mode (ATM). ATM is a variable-bandwidth, low-delay, packet-like switching and multiplexing technique. The multiplexed information flow is organized into blocks of a fixed size called cells.

Additional functionalities on top of the ATM layer, in the ATM Adaptation Layer (AAL), have to be provided to accommodate various applications. One of these applications is broadband signalling. In this report an AAL for signalling will be described. Report R34 (may 1990) of Study Group XVIII will be used as a starting point [1].

An introduction into B-ISDN is given in chapter 1. In chapter 2 the set of signalling protocols in the N-ISDN are described.

The N-ISDN will evolve during a period of time towards the B-ISDN. Therefore the protocol developed for the AAL for signalling have to remain 'upwards compatible'. The influence of this 'compatibility' and the influence of the underlying ATM layer on the AAL is described in chapter 3. The services and functions of the AAL for signalling are described in chapter 4. The structure and coding of the sublayers of the AAL are described in chapter 5. In chapter 6 the proposed AAL protocol and several functions are described.

Chapter 1

Broadband-ISDN

1.1 Introduction

During the nineties the ISDN will be implemented in several countries around the world. Because of the overall need for broadband services, which can not be supported by the ISDN, the ISDN will evolve towards a Broadband-ISDN (B-ISDN). The main feature of the B-ISDN is the support of a wide range of audio, video and data applications in the same network. A key element of service integration for a B-ISDN is the provision of a wide range of services to a broad variety of users utilizing a limited set of connection types and multipurpose user/network interfaces.
An overview will be given of the classification of the services in paragraph 1.2. The layered structure of B-ISDN will be explained in paragraph 1.3 and the characteristics of the transfer mode of the B-ISDN (Asynchronous Transfer Mode) are shown in paragraph 1.4.

1.2 Service Classification

This section describes the classification of broadband services and the definition of those service classes. Two main service categories are: interactive services and distributive services. The subdivision of these categories:

Interactive services:
  Conversational services
  Messaging services
  Retrieval services
Distribution services:
  Distribution services without user individual presentation control
  Distribution services with user individual presentation control

1.2.1 Conversational services

Conversational services in general provide the means for bidirectional communication with real-time (no store-and-forward) end-to-end information transfer from user to user or between user and host. The flow of the user information may be bidirectional symmetric, bidirectional asymmetric and in some specific cases (e.g. such as video surveillance), the flow of information may be unidirectional. The information is generated by the sending user or users, and is dedicated to one or more of the communication partners at the receiving site. Examples of broadband conversational services are videotelephony, video conference and high speed data transmission.

1.2.2 Messaging services

Messaging services offer user-to-user communication between individual users via storage units with store-and-forward, mailbox and/or message handling (e.g. information editing, processing and conversion) functions. Examples of broadband messaging services are message handling services and mail service for moving pictures, high resolution images and audio information.

1.2.3 Retrieval services

The user of retrieval services can retrieve information stored in information centers and in general provided for public use. This information will be sent to the user on his demand only. The information can be retrieved on an individual basis. Moreover, the time at which an information sequence is to start is under the control of the user. Examples are broadband retrieval services for film, high resolution image, audio information and archival information.

1.2.4 Distribution services without user individual presentation control

An important step in the evolution from the ISDN towards the B-ISDN is the integration of the CATVs and the telecommunication network. The broadcast services, which are now provided by a CATV, are an example of this type of distribution services. Distribution services provide a continuous flow of information which is distributed from a central source to a number of authorized receivers connected to the network. The user can access this flow of information without the ability to determine at which instant the distribution of a string of information will be started. The user cannot control the start and order of the presentation of the broadcast information. Examples are broadcast services for television and audio programs.

1.2.5 Distribution services with user individual presentation control

Services of these class also distribute information from a central source to a large number of users. However, the information is provided as a sequence of information entities (e.g. frames) with cyclical repetition. So, the user has the ability of individual access to the cyclical distributed information and can control start and order of presentation. Due to the cyclical repetition, the information entities selected by the user will always be presented from its beginning. An example of such a service is full channel broadcast videography.

1.3 Protocol Reference Model

The principles of layered communication are defined in Recommendation X.200: the reference model of Open Systems Interconnection (OSI) for CCITT applications. The OSI Reference Model is an abstract model and provides a conceptual an functional framework. It does not specify how the systems are to be implemented. The OSI Reference Model has seven layers, each with specific functions and offering defined services to the layer above and utilizing services of the layer below. This logical architecture should be applicable to the broadband network. An exact relationship between the lower layers of the OSI model and the lower layers of the B-ISDN Protocol Reference Model has not been defined and will probably not be defined also.

The Protocol Reference Model is shown in figure 1.1. It is composed of a user plane, a control plane and a management plane. Above the Physical Layer, the ATM Layer provides call transfer for all services and the AAL provides service-dependant functions to the layer above the AAL.

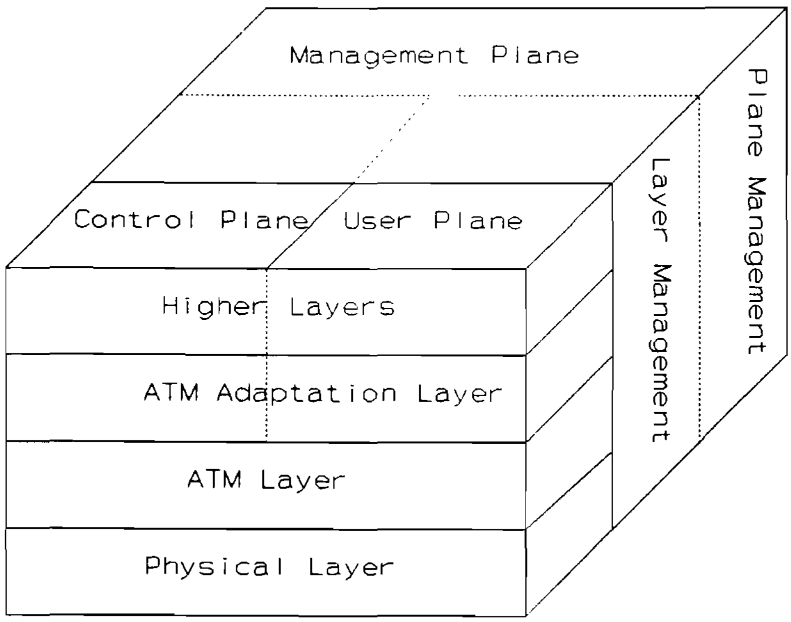


Figure 1.1: B-ISDN Protocol Reference Model [1]

1.3.1 Description of the planes

The User Plane, with its layered structure, provides for user information flow transfer, along with associated real-time controls (e.g. flow control, error recovery, etc.)

The Control Plane has also a layered structure and performs the call control and connection control functions; it deals with the signalling necessary to set up, modify and release calls and connections.

The Management Plane provides two types of functions, namely Layer Management and Plane Management:

- The Plane Management performs management functions related to a system as a whole and provides coordination between all the planes. Plane Management has no layered structure.
- Layer Management performs management functions relating to resources and parameters residing in its protocol entities. Layer Management handles the Operation and Maintenance (OAM) information flows specific to the layer concerned.

1.3.2 Functions of the individual layers of the B-ISDN PRM

This section gives a short description of the functions of the PRM layers. Table 1.1 illustrates the layers of the PRM, and identifies the functions of the Physical Layer, the ATM Layer and the AAL.

The Physical Layer consist of two sublayers. The Physical Medium sublayer (PM) includes only physical medium dependant functions like bit transfer, bit alignment, line coding and electrical-optical transformation. The Transmission Convergence (TC) sublayer performs all functions required to transform a flow of cells into a flow of data unit (e.g. bits) which can be transmitted and received over a physical medium.

Table 1.1: Functions of the B-ISDN in relation to the Protocol Reference Model [1]

Functions	Sublayer	Layer
Higher Layer Functions	Higher layers	
Convergence	CS	AAL
Segmentation and reassembly	SAR	AAL
Generic flow control Cell header generation/extraction Cell VPI/VCI translation Cell multiplex and demultiplex	ATM	
Cell rate decoupling HEC header sequence generation/verification Cell delineation Transmission frame adaptation Transmission frame generation/recovery	TC	Physical Layer
Bit timing Physical medium	PM	Physical Layer

1.4 ATM Layer

1.4.1 Introduction

A user of the B-ISDN can use multiple services. Therefore an important issue within the B-ISDN is the availability of a flexible transfer mode. ATM is the transfer mode solution, chosen by the CCITT Study Group XVIII, for implementing a B-ISDN. It influences the standardization of digital hierarchies, multiplexing structures, switching and interfaces for broadband signals. ATM is a packet-oriented transfer mode which uses asynchronous time division multiplexing techniques. The multiplexed information flow is organized into blocks of a fixed size called cells. A cell consists of an information field and a header. The primary role of the header is to identify cells belonging to the same virtual channel within the asynchronous time division multiplex. ATM is a connection-oriented technique. Virtual channel identifiers are assigned to each logical connection when required and released when no longer needed. Transfer capacity is assigned by negotiation and is based on the source requirements and the available capacity. Cell sequence integrity on a virtual channel connection is preserved by the ATM layer. In general, signalling and user information are carried on separate ATM Layer connections.

1.4.2 ATM Layer Connections

The transport functions of the ATM layer are divided into two levels, the Virtual Channel (VC) level and the Virtual Path (VP) level. Each ATM cell contains a label in its header to identify the Virtual Channel to which the cell belongs. This label consists of two parts: a Virtual Channel Identifier (VCI) and a Virtual Path Identifier (VPI). A Virtual Channel is only fully identified at an interface by both VPI an VCI values. Virtual Channel links are concatenated to form a Virtual Channel Connection (VCC). A Virtual Path is a bundle of virtual channel links. Virtual path links are concatenated to form a Virtual Path Connection (VPC).

Different Virtual Path links, which are multiplexed at the ATM layer into the same physical layer connection are distinguished by the VPI. The different Virtual Channel links in a Virtual Path Connection are distinguished by the VCI as indicated in figure 1.2.

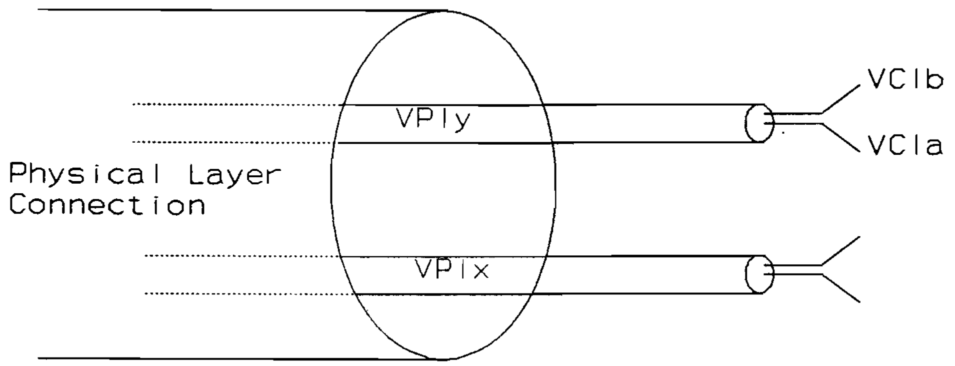


Figure 1.2: ATM layer connection identifiers [1]

A specific value of VCI is assigned each time a VC is switched in the network. A VC link is a unidirectional capability for the transport of ATM cells between two ATM entities where the VCI value is translated. Important topics for a VCC/VPC are:

- Each VCC/VPC provides a *Quality of Service* specified by parameters such as cell loss ratio and cell delay variation;
- A VCC/VPC can be provided on a *switched or (semi-)permanent* basis;
- *Cell sequence* is preserved within a VCC/VPC;
- *Traffic parameters* shall be (re-)negotiated between a user and a network and the network will ensure that these parameters are not violated.

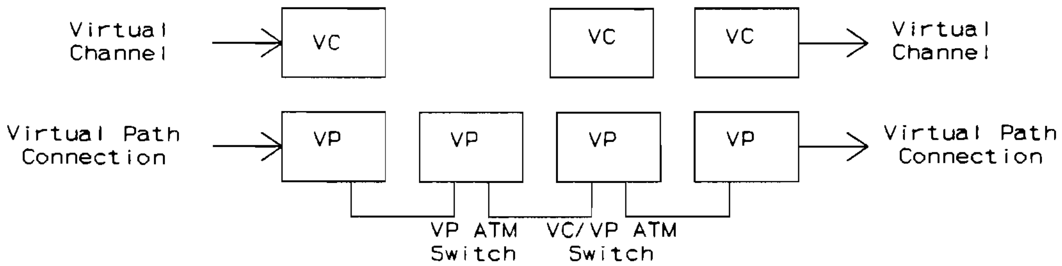


Figure 1.3: Types of ATM layer connections [1]

The basic ATM routing entity for switched services is the Virtual Channel. It is handled in VC multiplexers/demultiplexers and switches. VCs are aggregated in VPCs which are routed as such through VP multiplexers/ demultiplexers and VP switches.

1.4.3 Cell Structure

The ATM cell consists of a header and an information field. The header has a length of five octets and the information field consists of 48 octets.

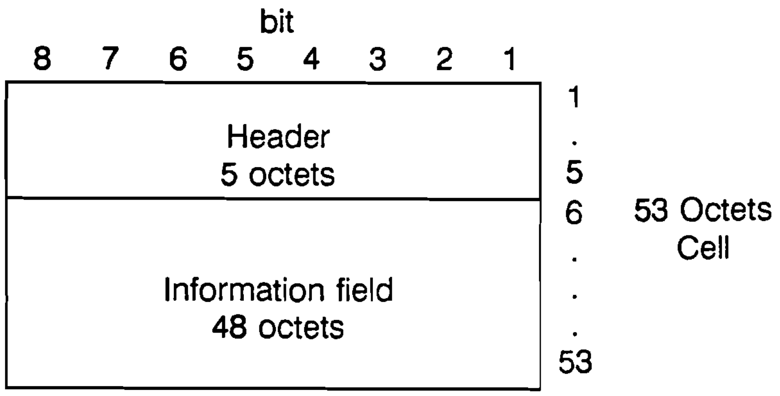
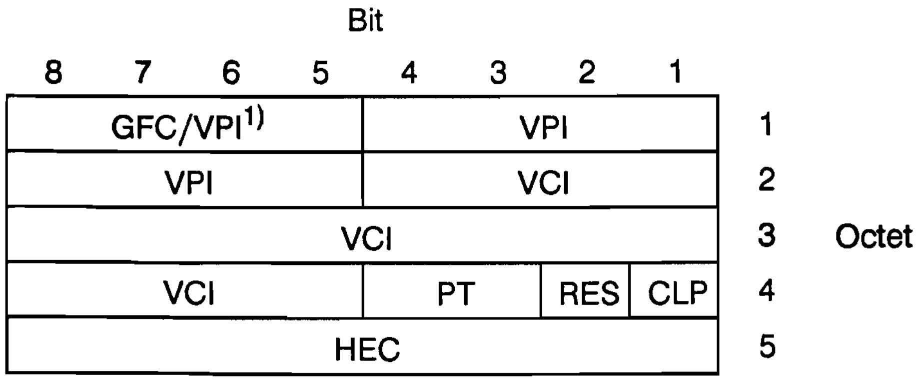


Figure 1.4: Cell structure at the UNI/NNI [1]

Two different cell header coding schemes are adopted according to the interface being considered, i.e. the User-Network Interface (UNI) or the Network-Node Interface (NNI). The structure of the header at the UNI/NNI is shown in figure 1.5.

The *Generic Flow Control* (GFC) field contains 4 bits. The GFC mechanism is used to control traffic flow in order to alleviate short-term overload conditions that may occur.

24 and 28 bits are available for routing at the UNI respectively the NNI. The *Virtual Path Identifier* contains 8 bits at the UNI and 12 bits at the NNI, while the *Virtual Channel Identifier* contains 16 bits at both the UNI and the NNI. Several pre-assigned combinations of VCI and VPI values are standardized for the UNI and the NNI.



1) UNI: GFC / NNI: VPI

Figure 1.5: Header structure at UNI/NNI [1]

The field *Payload Type* (PT) contains 2 bits and will be used for identification of the information field. It is used to provide an indication of whether the cell payload contains user information or network information. The default value for user information is 00.

The *Cell Loss Priority* (CLP) field gives each cell a priority. If the bit is set, the cell is subject to discard, depending on network conditions. If the CLP is not set, the cell has a higher priority. The rate of higher priority cells will be determined at call establishment and may be subsequently renegotiated. The CLP mechanism would not normally be used for Constant Bit Rate (CBR) services.

The *Header Error Control* (HEC) field consists of 8 bits. The HEC covers the entire cell header. The code used for this function is can correct single bit errors and can detect multiple-bit errors.

The *Reserved* (RES) field (1 bit) is for further enhancement of existing cell header functions or for standardized functions not yet specified.

1.5 Conclusions

A 'next step' in the world of telecommunications will be the Broadband-ISDN. B-ISDN will support a wide range of services. These services can be divided into two main service categories: interactive services and distribution services.

The objective of the B-ISDN is to be able to provide all possible service mixes by using the same transmission technique. The transfer mode of the B-ISDN is the Asynchronous Transfer Mode (ATM). ATM can be defined as a packet-oriented switching and multiplexing technique which is expected to offer user satisfactory quality of service and high bandwidth utilization through statistical multiplexing and high flexibility.

In the ATM layer the multiplexed information flow is organized into blocks of a fixed size, called cells, consisting of an information field of 48 octets and a header of 5 octets. The primary role of the header is to identify the cells belonging to the same virtual channel. In the ATM cell header one bit is defined as Cell Loss Priority bit (CLP). For ATM carrying signalling traffic it is necessary that the CLP bit is not set to achieve a low cell loss rate.

Above the ATM layer the AAL provides service dependant functions to the layer above the AAL. This AAL is separated into two sublayers: the Convergence Sublayer (CS) and the sublayer Segmentation and Reassembly (SAR).

Chapter 2

Signalling in ISDN

2.1 Introduction

Signalling can be defined as "the exchange of information specially concerned with the establishment and control of connections and with network management, in a telecommunications network" [2]. Signalling constitutes the command/control infrastructure of the modern telecommunications networks. It is something "analogous to the central nervous system of a living organism, something to coordinate the functions while remaining completely separate from the organism's other parts and not actually performing their function".

Signalling has traditionally consisted of supervisory functions [3] (e.g., on-hook/ off-hook to indicate idle or busy status), addressing (e.g., called number), and providing call information (e.g., dial tone and busy signals). Common control switching systems, like the No. 5 Cross Bar, introduced Single-Frequency (SF) / Multi-Frequency (MF) signalling techniques. With the introduction of electronic processors in switching systems came the possibility of providing common channel signalling. This is an out-of-band signalling method in which a common data channel to convey signalling information related to a number of trunks.

The advent of the ISDN has accelerated the pace of development and deployment of signalling systems to support an ever increasing set of "intelligent network" services on a world wide basis.

Signalling information can be exchanged user to user, network-node to network-node and user to network-node as shown in figure 2.1.

In N-ISDN there are two signalling protocols (access signalling and network signalling). The current set of protocol standards for N-ISDN access signalling is known as the Digital Subscriber Signalling No. 1 (DSS 1). The current set of protocol standards for N-ISDN network signalling is known as the Signalling System No. 7 (SS7).

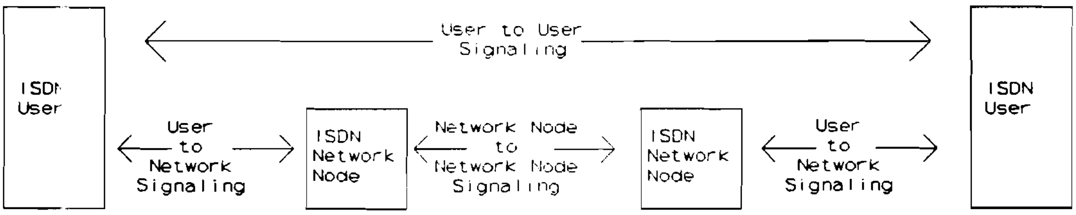


Figure 2.1: Signaling information exchanges [2]

These protocols are the basis for providing the real-time control and operation of today's distributed telecommunications network by supporting the exchange of information between intelligent elements such as terminals, switches and databases within and between private and public networks.

2.2 Signalling System No.7

Figure 2.2 shows the architecture of the SS7 protocol. The SS7 component that correspond to the first three layers of the OSI reference model is called the Network Service Part (NSP). It consist of the Message Transfer Part (MTP) and the Signalling Connection Control Part (SCCP).

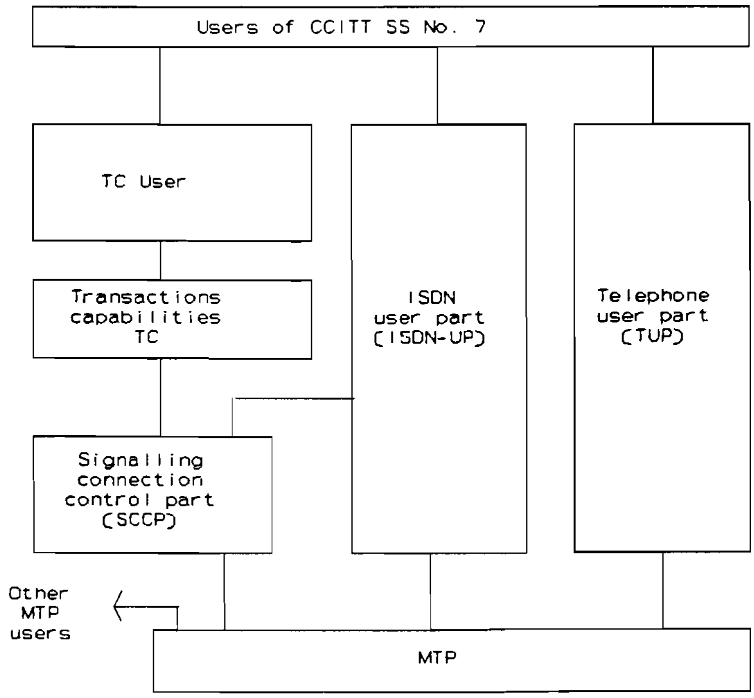


Figure 2.2: Architecture of CCITT SS No.7 [3]

MTP provides a connectionless message transfer system that enables signalling information to be transferred across the network to its desired destination. Functions are included in MTP that allow system failures to occur in the network without adversely affecting the transfer of signalling information. MTP was developed before SCCP and it was tailored to the real-time needs of telephony applications. Later, it became clear that there were other applications that would need additional network services like an expanding addressing capability and connection-oriented message transfer. SCCP was developed to satisfy this need.

2.2.1 Message Transfer Part

The overall purpose of the MTP is to provide a reliable transfer and delivery of signalling information across the signalling network and to have the ability to react and take necessary actions in response to system and network failures to ensure that reliable transfer is maintained. Figure 2.3 shows the structure of the signalling system functions. In relation to the development of an AAL for signalling the level 2 and 3 functions are the most important and will therefore described here.

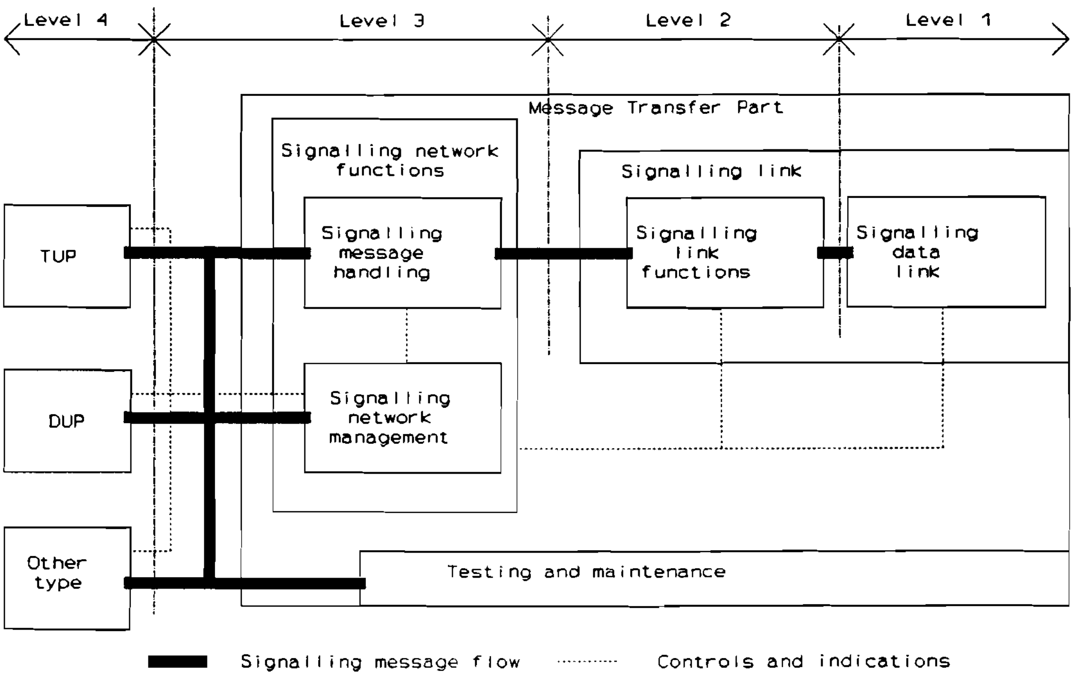


Figure 2.3: General structure of signaling system functions [3]

Signalling link functions

A signalling data link is a bidirectional transmission path for signalling, consisting of two data channels operate together in opposite directions at the same data rate. It fully complies with the OSI's definition of the physical layer. For digital signalling data links, the recommended bit rate for the CCITT international standard is 64 kb/s.

The signalling link functions correspond to the OSI's data link layer. Together with a signalling data link, the signalling link functions provide a signalling link for the reliable transfer of signalling messages between two directly connected signalling points. Signalling messages are transferred over the signalling link in a variable length message called *signal units*. The signalling link functions are:

- signal unit delimitation/alignment
- error correction/detection
- initial alignment
- signalling link error monitoring
- flow control

There are three types of signal units, differentiated by the length indicator contained in all signal units, as shown in figure 2.4.

F	CK	SIF	SIO		LI	FIB	FSN	BIB	BSN	F
8	16	8n, n>=2	8	2	6	1	7	1	7	8

a) Basic format of a message signal unit (MSU)

F	CK	SF		LI	FIB	FSN	BIB	BSN	F
8	16	8 or 16	2	6	1	7	1	7	8

b) Format of a link status signal unit (LSSU)

F	CK		LI	FIB	FSN	BIB	BSN	F
8	16	2	6	1	7	1	7	8

LI: 0 FISU
1 or 2 LSSU
> 2 MSU

c) Format of a fill-in signal unit (FISU)

F	Flag	FIB	Forward indicator bit
CK	Check bits	FSN	Forward sequence number
SIF	Signaling information field	BIB	Backward indicator bit
SIO	Service information octet	BSN	Backward sequence number
LI	Length indicator	SF	Status field

Figure 2.4: Signal units format [3]

The length indicator is used to indicate the number of octets following the length indicator octet and preceding the check bits. The indicator bits together with the sequence numbers are used in the basic error control method to perform the signal unit sequence control and acknowledgement functions. The error detection function is performed by means of the standard CCITT 16-bit Cyclic Redundancy Check (CRC) checksum. FISUs are sent when there is no message traffic so that faulty links can be quickly detected and removed from service even when traffic is low.

Two forms of error correction are specified in the signalling link procedures. They are the Basic method and the Preventive Cyclic Retransmission (PCR) method. In both methods only errored MSUs and LSSUs are corrected (errors in FISUs are detected but not corrected).

The basic method of error correction is a noncompelled positive/negative acknowledgement retransmission error correction system. It uses the 'go-back-N' technique of retransmission used in many other protocols. If a negative acknowledgement is received, the transmitting terminal stops sending new MSUs and rolls back to the MSU received in error and it retransmits everything from that point before resuming transmission of new MSUs. Positive acknowledgments are used to indicate correct reception of MSUs and as an indication that the positively acknowledged buffered MSUs can be discarded at the transmitting end.

The preventive cyclic retransmission method is non-compelled, positive acknowledgement, retransmission error correction system. A MSU must be retained at the transmitting signalling link terminal until a positive acknowledgement arrives for that MSU. When there are no MSUs to be sent, all MSUs not positively acknowledged are retransmitted cyclically. When the number of unacknowledged MSUs exceeds certain thresholds, it is an indication that error correction is not getting done by cyclic retransmission and then a forced retransmission procedure is invoked until the number of unacknowledged messages is below the threshold value.

Two signalling error rate monitor functions are provided; one which is used while a link is in service and which provides one of the criteria for taking the link out of service and one which is used while a link is in the providing state of the initial alignment procedure. These are called the signal unit error rate monitor and the alignment rate error rate monitor.

The flow control procedure is initiated when congestion is detected at the receiving end at the signalling link. The congested receiving end notifies the transmitting end of its congestion with an LSSU indicating busy and withholds acknowledgements of all incoming signal units. The action stops the transmitting end from failing the link due to a time out on acknowledgements. However, if the congestion condition lasts too long, the transmitting end will fail the link.

Signalling network functions

The signalling network functions correspond to the lower half of the OSI's network layer and they provide the functions and procedures for the transfer of messages between signalling points. As shown in figure 2.3, the signalling network functions can be divided into two basic categories, namely signalling message handling and signalling network management.

Signalling message handling

The purpose of the signalling message handling functions is to ensure that the signalling messages originated by a particular User Part at a signalling point (originating point) are delivered to the same User Part at the destination point indicated by the sending User Part. Signalling message handling consist of message routing, discrimination and distribution functions.

Signalling network management

The purpose of the signalling network management functions is to provide reconfiguration of the signalling network in the case of signalling link or signalling point failures and to control traffic in the case of congestion. When a failure occurs, the reconfiguration is carried out so that messages are not lost, duplicated or put out of sequence and that the message delays do not become excessive. Signalling network management consist of three functions: signalling traffic management, signalling route management and signalling link management.

2.2.2 SCCP

SCCP enhances the services of the MTP to provide the functional equivalent of OSI's Network layer. SCCP provides four classes of service, two connectionless and two connection oriented. The four classes are:

- Class 0: Basic connectionless class
- Class 1: Sequenced (MTP) connectionless class
- Class 2: Basic connection-oriented class
- Class 3: Flow control connection-oriented class

Class 0 is a pure connectionless: the Network Service Data Units (NSDUs) are transported independently and may be delivered out of sequence.

In Class 1 the SCCP assures that a particular message stream of NSDUs will be delivered in sequence by sending all messages in the stream over the same link.

In Class 2, a bidirectional transfer of NSDUs is performed by setting up a temporary or permanent signalling connection. This service class provides in sequence delivery of messages and also a segmentation and reassembly capability.

In Class 3, the capabilities of Class 2 are provided with the addition of flow control. Also the detection of message loss and mis-sequencing is provided. In the event of lost or mis-sequenced messages, the signalling connection is reset and notification is given to the higher layers

2.3 Digital Subscriber Signalling No. 1

2.3.1 Data Link Layer

This section describes the Link Access Procedure on the D-Channel, LAPD [4]. LAPD is a protocol that operates at the data link layer of the OSI architecture. The purpose of LAPD is to convey information between layer 3 entities across the ISDN user-network interface using the D-channel. LAPD will support multiple terminal installations at the UNI and multiple layer 3 entities. All data link layer messages are transmitted in frames which are delimited by flags. LAPD includes functions for:

- the provision of one or more data link connections on a D-channel. Discrimination between the data link connections is by means of a Data Link Connection Identifier (DLCI) contained in each frame;
- frame delimiting, alignment and transparency, allowing recognition of a sequence of bits transmitted over a D-channel as a frame;
- sequence control, to maintain the sequential order of frames across a data link connection;
- detection of transmission, format and operational errors on a data link connection;
- recovery from detected transmission, format and operational errors;
- notification to the management entity of unrecoverable errors;
- flow control.

Data link layer functions provide the means for information transfer between multiple combinations of data link connection endpoints. The information transfer may be via point-to-point data link connections or via broadcast data link connections. In the case of point-to-point information transfer, a frame is directed to a single endpoint, while in the case of broadcast information transfer, a frame is directed to one or more endpoints.

All data link layer peer-to-peer exchanges are in frames conforming to one of the formats shown in figure 2.5.

8 7 6 5 4 3 2 1	8 7 6 5 4 3 2 1	
Flag 0 1 1 1 1 1 1 0	Flag 0 1 1 1 1 1 1 0	Octet 1
Address (high order octet)	Address (high order octet)	2
Address (low order octet)	Address (low order octet)	3
Control	Control	4
	Control	
Control	Information	
FCS (first octet)	FCS (first octet)	N - 2
FCS (second octet)	FCS (second octet)	N - 1
Flag 0 1 1 1 1 1 1 0	Flag 0 1 1 1 1 1 1 0	N

Figure 2.5: Frame formats [4]

All frames start with the flag sequence consisting of one 0 bit followed by six contiguous 1 bits and one 0 bit. The address field consists of two octets. The address field identifies the intended receiver of a command frame and the transmitter of a response frame. The format of the address field is shown in figure 2.6.

8 7 6 5 4 3	2	1	
SAPI	C/R	EA 0	Octet 2
TEI		EA 1	3

EA = Address field extension bit
C/R = Command/response field bit
SAPI = Service access point identifier
TEI = Terminal endpoint identifier

Figure 2.6: Address field format [4]

It contains the address field extension bits, a command/response indication bit, a data link layer Service Access Point Identifier subfield and a Terminal Endpoint Identifier.

A data link connection is identified by a Data Link Connection Identifier (DLCI) carried in the address field of each frame. The DLCI consist of two elements. The SAPI and the Terminal Endpoint Identifier (TEI). The SAPI is used to identify the service access point on the network side or the user side of the UNI. The TEI is used to identify a specific connection endpoint within a service access point. The DLCI is a pure data link layer concept. The allocated SAPI and TEI values are shown in figure 2.7.

SAPI Value	Related layer 3 or management entity
0	Call control procedures
1	Packet mode communications using Q.931 call control procedures
16	Packet communication conforming to X.25 Level 3 procedures
63	Layer 2 Management procedures
All Others	Reserved for future standardization

TEI Value	User Type
0-63	Non-automatic TEI assignment user equipment
64-126	Automatic TEI assignment user equipment
127	Broadcast

Figure 2.7: SAPI and TEI allocation [4]

The TEI for a point-to-point data link connection may be associated with a single Terminal Equipment (TE). A TE may contain one or more TEIs used for point-to-point data transfer.

The address field range is extended by reserving the first transmit bit of the address field octets to indicate the final octet of the address fie

The C/R bit identifies a frame as either a command or a response. The user sends commands with the C/R bit set to 0 and responses with the C/R bit set to 1. The network side does the opposite.

The control field identifies the type of frame which will be either a command or response. The control field will contain sequence numbers, where applicable. Three types of control field formats are specified: numbered information transfer (I format), supervisory functions (S format) and unnumbered information transfers and control functions (U format).

The information field, when present, follows the control field and precedes the frame check sequence and consists of layer 3 information.

The FCS is used to detect errors. When a data link endpoint receives a frame with an incorrect FCS it will be discarded and the endpoint will ask for retransmission. The data link endpoint can detect lost frames by the numbering in the control field.

2.3.2 Network Layer

The layer 3 protocol provides the means to establish, maintain and terminate network connections across an ISDN between communicating application entities. There are two categories of functions performed at layer 3 and services provided by layer 3 in the establishment of network connections. The first category contains these functions which directly control the connection establishment. The second category contains those functions relating to the transport of messages additional to the functions provided by the data link layer.

Functions performed by layer 3 include the following:

a) processing of primitives for communicating with the data link layer;
b) generation and interpretation of layer 3 messages for peer-level communication;
c) administration of timers and logical entities used in the call control procedures;
d) administration of access resources including B-channels and packet-layer logical channels;
e) checking to ensure that services provided are consistent with user requirements;

Within this layer 3 protocol, every message consists of the following parts:

a) protocol discriminator
b) call reference
c) message type
d) other information elements, as required.

Information elements a), b) and c) are common to all the messages and are always present, while information element d) is specific to each message type. The structure of the message is shown in figure 2.8.

Bit 8	7	6	5	4	3	2	1	Octet
Protocol discriminator								1
0	0	0	0	Length of call ref. val.				2
Call reference value								3
0	Message type							etc.
Other information elements as required								

Figure 2.8: Structure of layer 3 message [4]

The purpose of the protocol discriminator is to distinguish messages for user-network call control from other messages. The purpose of the call reference is to identify the call or registration/cancellation request at the local user-network interface to which the particular message applies. The call reference does not have end-to-end significance across ISDNs. Call reference values are assigned by the originating side of the interface for a call. The purpose of the message type is to identify the function of the message being sent.

Two categories of information elements are defined:
- single octet information elements
- variable length information elements.

The format of the variable length information element is shown in figure 2.9.

8	7	6	5	4	3	2	1	Octet
0	Information element identifier							1
Length of contents of information element (octets)								2
Contents of information element								3 etc.

Figure 2.9: Variable length information element format [4]

A particular message may contain more information than a particular (user or network) equipment needs or can understand. So, a particular equipment can ignore a certain information element and the number of octets to ignore, will be dependant of the value octet 2.

2.4 Conclusions

In an ISDN there are two signalling protocols (access signalling and network signalling).

The set of protocol standards for ISDN access signalling is known as the Digital Subscriber Signalling No. 1 (DSS1). DSS1 specifies in Q.921 LAPD, a protocol that operates at the data link layer of the OSI architecture. DSS1 also specifies a layer 3 protocol that is designed to effect the establishment and control of circuit-switched and packet-switched connections.

The current set of protocol standards for ISDN network signalling is known as the Signalling System No.7 (SS7). The most important part of SS7, in relation to the development of an AAL for signalling, is the Message Transfer Part (MTP). MTP provides a connectionless message transfer system that enables signalling information to be transferred across the network to its desired destination.

DSS1 and SS7 are converging toward a single control protocol whose architecture will be based upon the OSI Reference Model and its protocols [6].

Chapter 3

Towards an AAL for signalling

3.1 Introduction

In the first chapter of this document an introduction was given about B-ISDN. The existing signalling protocols for an ISDN (SS no. 7, DSS1) were explained in the second chapter.

B-ISDN must be compatible with and evolve from existing networks, so that these networks and their services can be interworked easily. The term compatibility used in this way is described in paragraph 3.2.

Before developing an AAL for signalling it is necessary to investigate the performance of the existing layer 2 protocols and the performance of the future ATM layer.

The performance of the layer 2 signalling protocols of ISDN and SS7 are decribed in paragraph 3.3.

The B-ISDN will be mainly based on optical fiber. Fiber has better error characteristics than twisted pair and coaxial cable. In paragraph 3.4 the performance of the physical layer and the ATM layer will be investigated.

3.2 Compatibility

3.2.1 UNI and NNI

In the ISDN there are two different signalling protocols (access signalling and network signalling). The ISDN access signalling protocol is known as the Digital Subscriber Signalling No. 1 (DSS 1). The ISDN network signalling protocol is known as the Signalling System No. 7 (SS7). The extended requirements of signalling at the NNI may not be required at the UNI, however selected subsets of a common AAL protocol should be used rather than devising separate AAL protocols. Therefore it is important to develop an AAL for signalling that is common for the UNI and the NNI. This approach is consistent with long term signalling requirements for unified UNI/NNI signalling.

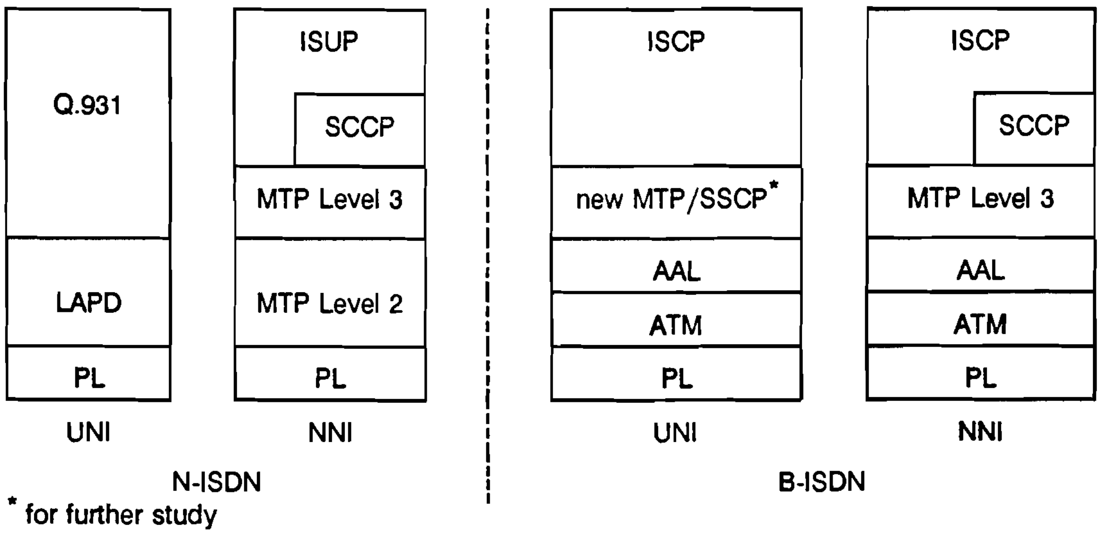


Figure 3.1: Control protocol stacks of (B-)ISDN

Figure 3.1 shows the protocolstacks in the control plane of the (B-)ISDN.

3.2.2 Upward compatibility

The AAL for signalling provides services to the network layer. Before developing such AAL it is important to investigate the performance of:

- the Data Link Layer (Q.921);
- the Message Transfer Part level 2;
- the ATM layer.

As shown in figure 3.1 is it possible to look at the AAL from the layer above and from the underlying ATM layer. First the Quality of Service of the existing layer 2 protocols have to be investigated. If the performance of the ATM layer is known, the AAL has to be developed in a way that at least the same Quality of Service can be provided to the layer 3 protocols. Until now the terms Quality of Service and Network Performance are used for the same subject. However, in I.350 a distinction is made between these two terms. The QOS is defined as follows: "Collective effect of service performance which determine the degree of satisfaction of a user of the services". On the other hand, the NP parameters determine the (user observed) QOS, but they do not necessarily describe that quality in a way that is meaningful to users. Parameters which are important to the transport of signalling messages are:

- error performance parameters
- message delay parameters

Figure 3.2 shows the Service Access Points which are important for developing an AAL for signalling.

The services provided at the SAPs 1 and 2 in figure 3.2 have to be compared with the future ATM service at SAP 3. The term compatibility can be used in this paragraph in a way that the future AAL for signalling have to be compatible with the currently existing layer 3 layers.

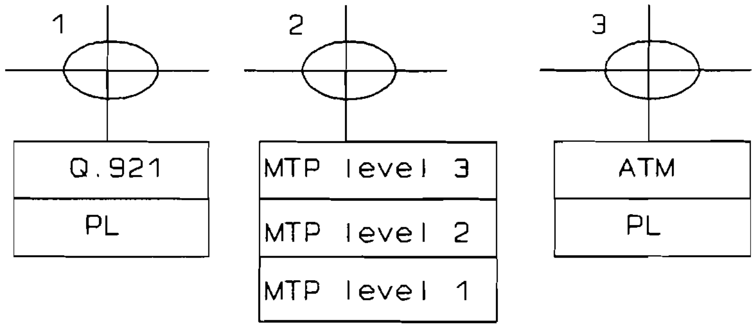


Figure 3.2: Important SAPs for developing an AAL for signalling

3.3 Performance of layer 2 protocols

3.3.1 MTP level 2

The Quality of Service provided by the Message Transfer Part level 2 is not specified in the recommendations of the Blue Book. The performance of the complete MTP is given in Q.706 [7]. The parameters of the signalling performance are divided into two subjects:

- error performance
- message transfer time

The transport of messages by the MTP is protected against transmission errors. This is done by check bits, length indicators, sequence numbers and error correction. The following conditions are therefore guaranteed by the MTP:

- Not more than one in 10^{10} of all signal unit errors will be undetected;
- Not more than one in 10^{7} messages will be lost;
- Not more than one in 10^{10} messages will be delivered out-of-sequence to the User Parts (including duplication).

During the transport of messages from one UP to another UP the MTP introduces a message transfer time. The overall message transfer time depends on five compounds:

T_{ms}	MTP sending function
T_{cs}	STP function
T_{mr}	MTP receiving function
T_p	Signalling data link propagation time
T_q	Queuing delay

Figure 3.3 shows a functional diagram of four components.

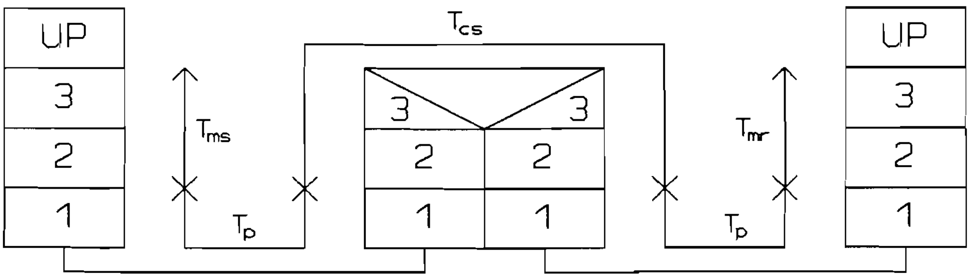


Figure 3.3: Message Transfer Delay Compounds of MTP [7]

Queuing delays are introduced by the MTP because the MTP has to transmit signal units from different User Parts. Another cause of delay is introduced in the case of disturbances when error correction (retransmission) affects the overall transfer time of a signal unit. Another parameter which affects the message transfer time is the traffic loading. In figure 3.4 the queuing delay of a MSU in the MTP in shown in the case of disturbances and the basic correction method. Figure 3.4 shows that at a traffic loading a < 0,3 the message transfer time is mainly determined by the time T_L (signalling loop propagation time). The mean total queuing delay, excluding the propagation time, is less than 2 ms at a traffic loading of 0,5 Erlang. The queuing delay compound will be greater when the PCR method is used for error correction, but also then the propagation time will be the dominant factor of the message transfer time.

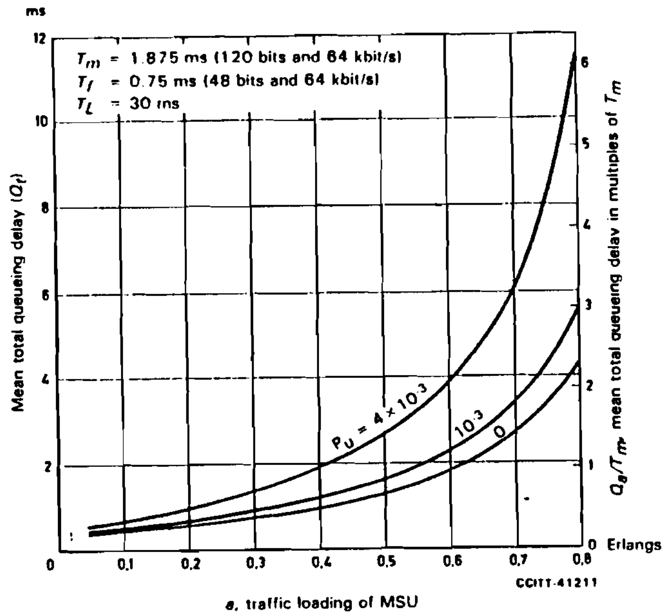


Figure 3.4: Mean total queuing delay of each channel of traffic [7]

3.3.2 Q.921: Data Link Layer DSS1

In recommendation Q.921 nothing is stated about the performance of the data link layer. General aspects of Quality of Service and Network Performance in an ISDN are listed in Recommendation I.350. In G.821 [8] there are descriptions of network performance parameters, that deal with the reliability of an ISDN network. The error performance parameters are defined as follows:

> *"The percentage of averaging periods each of time interval To during which the bit error ratio (BER) exceeds a threshold value. The percentage is assessed over a much longer time interval T_L".*

The following BERs and intervals are used in the statements of objectives:

a) a BER of less than 10^{-6} for To = 1 minute
b) a BER of less than 10^{-3} for To = 1 second
c) zero errors for To = 1 second

The performance objectives for an international ISDN connection are shown in table 3.1.

Table 3.1: Error performance objectives for international ISDN connections [8]

Performance classification	Objective
Degraded minutes	Fewer than 10% of one-minute intervals to have a BER worse than 10^{-6}
Severely errored seconds	Fewer than 0.2% of one-second intervals to have a BER worse than 10^{-3}
Errored seconds	Fewer than 8% of one-second intervals to have any errors (equivalent to 92% error-free seconds)

The connection fails to satisfy the objective if any of the requirements is not met. Table 3.1 applies to a 64 kbit/s circuit switched connection.

3.4 ATM Characteristics

3.4.1 Fiber Optic Links

The B-ISDN will be based on fiber links. Therefore it is necessary to investigate the error characteristics of a fiber optic system. In [9] observations of error characteristics of fiber optic transmission systems are reported. A fiber optic system, installed in the field to carry future service traffic and capable of carrying many DS-3 rate (45 Mb/s) channels per fiber was monitored over a 94 day period. In figure 3.5 a histogram is shown of the relative numbers of errors per errored second over a period of 92 days.

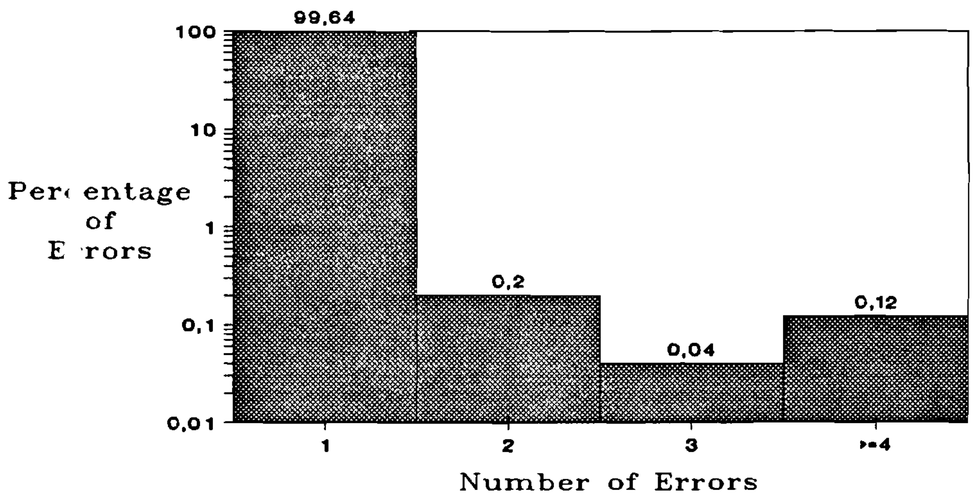


Figure 3.5: Percent of Errors per Errored Second [9]

Figure 3.5 shows that most of the ES have one bit error. However in long haul transmission systems there are also burst errors caused predominantly by protection switching. A failed repeater will cause transmission of messages to be switched onto a protection line. During the switching and reframing process the circuit will be essentially open resulting in a high error rate for a small fraction of a second.

It is reported in [10] that a 950 mile long, high capacity digital fiber transmission system experienced an average of 0.24 protection switches per day. The duration of the switching and reframing process is between 20 and 40 milliseconds and for this duration the Bit Error Rate (BER) is approximately 0.5. During a burst all the cells will be lost. The number of cells transmitted during 40 ms is 6 for a 64kb/s link and about 14,000 for a 150 Mb/s link.

So, in fiber optic systems:

- the background errors are predominantly single bit errors;
- the major source of burst errors is protection switching, with burst errors lasting 20 - 40 ms.

3.4.2 Cell Errors and Cell Delay

In this paragraph the impact of errors, described in paragraph 3.4.1, on ATM cells will be shown. Transmission bit errors can affect the header field and/or the information field of an ATM cell. In figure 3.6 is shown the probability of the number of bit errors in the header or the information field dependant on the BER.

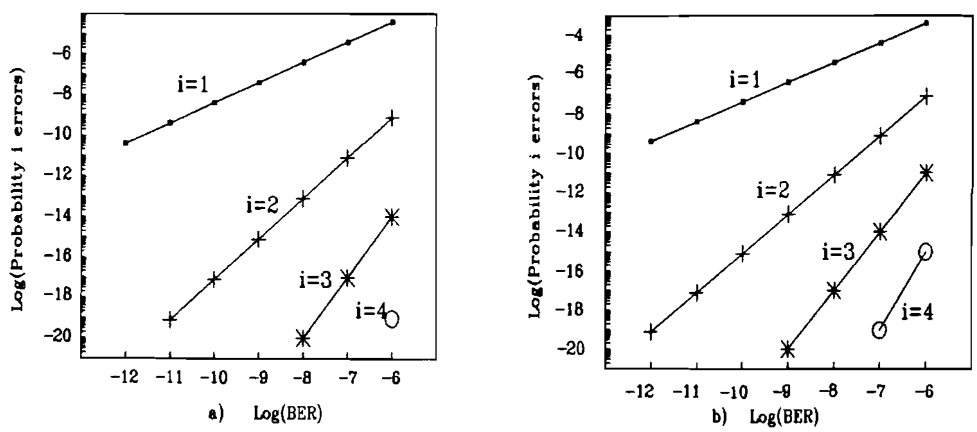


Figure 3.6: a) Probability of an errored header vs BER; b) Probability of an errored info field vs BER

The ATM layer detects and/or corrects the bit errors in the cell header by a Header Error Control (HEC) mechanism. The detection and correction of the errors in the information field will be done in the higher layers.

In an ATM cell the Header Error Control covers the entire cell header. The code used for this function is capable either of;

- single-bit error correction;
- multiple-bit error correction.

In figure 3.7 the consequences are shown of errors in an ATM cell header for the HEC process.

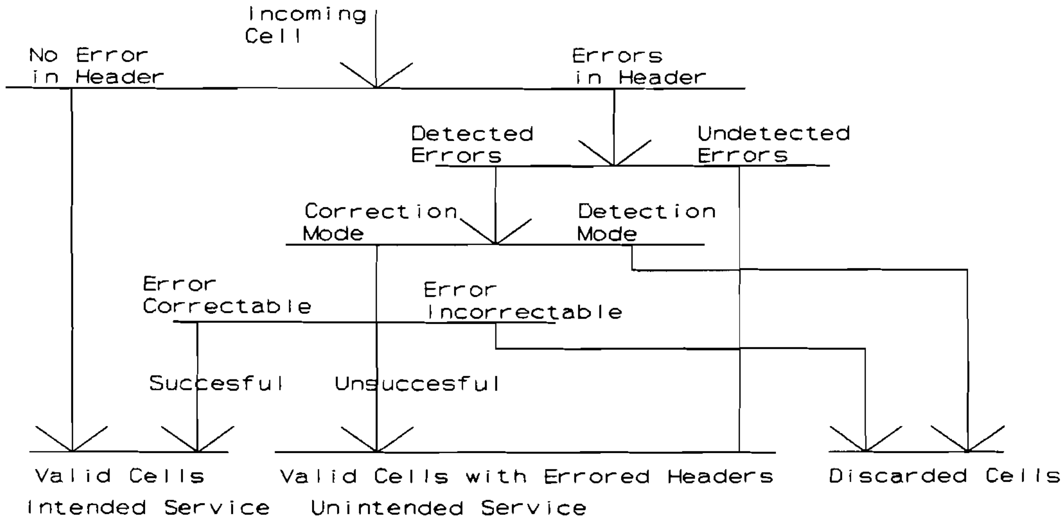


Figure 3.7: Header Error Control process [1]

In figure 3.8 is shown how random bit errors impact the probability of occurrence of discarded cells and valid cells with errored headers.

In general a connection is divided into sections. Each section represents the transmission between user and local exchange or between exchanges. After each route section one of the following 5 cases can occur. The cell is either:

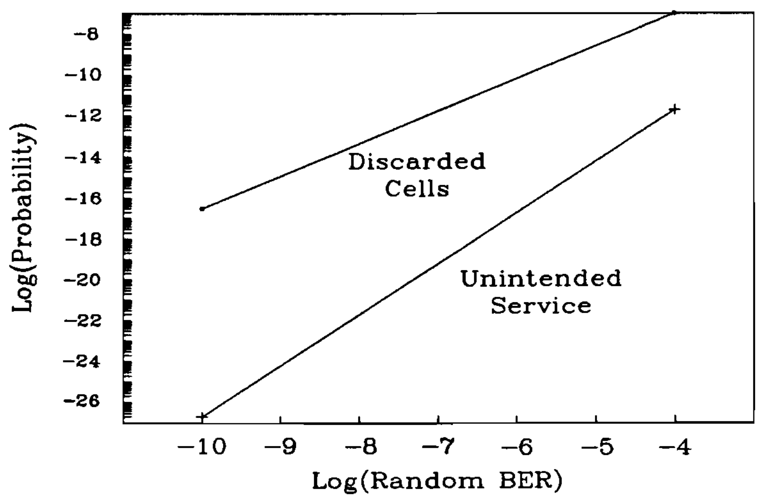


Figure 3.8: Unintended service and discarded cell probability [1]

- correctly routed;
- misrouted on the same link with a foreign VCI;
- misrouted onto a wrong link with a foreign VCI;
- the cell got lost, because the VCI is affected by bit errors and did not match a used VCI;
- the cell got lost, because the VCI is affected by bit errors and does not match a valid header.

In [11] the impact of bit errors and the number of route sections on the Cell Loss Rate (CLR) and the Cell Insertion Ratio (CIR) is reported. This is shown in figure 3.9.

Figure 3.9 is based on the assumptions:

- all route sections provide a BER of 10^{-6} at the same time
- all the bit errors occur uncorrelated
- there are less than 10 sections used in a connection

The result in figure 3.9 is, the CLR due to bit errors is less than 10^{-9}. The CIR of less than 10^{-14} can be negelected.

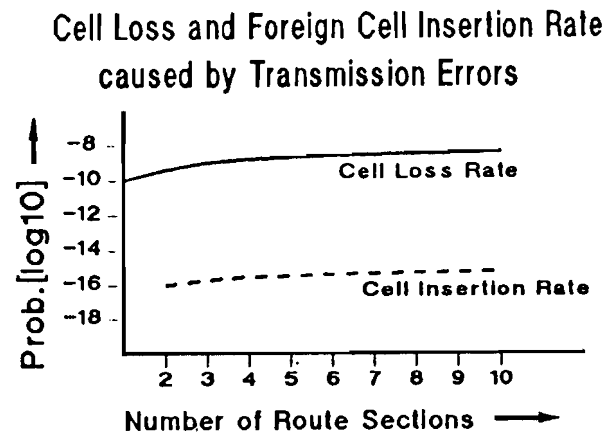


Figure 3.9: Cell Loss and Cell Insertion Rate caused by transmission errors [11]

The end-to-end delay T_a of ATM cells depends on several compounds [36]:

T_e due to CODEC
T_c cell assembly/disassembly
T_n transfer delay (waiting times)
- transit nodes
- multiplexing
- header error correction/detection

T_p link propagation delay
T_d cell buffering

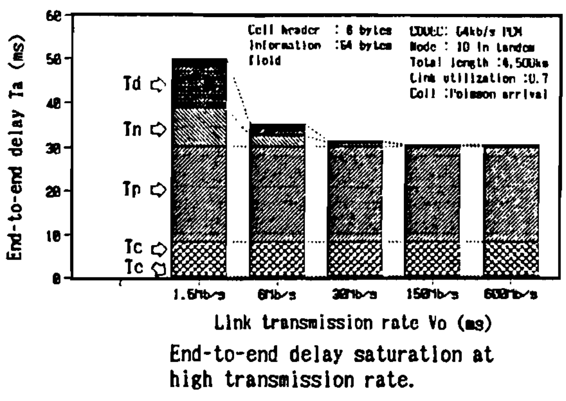


Figure 3.10: Delay ATM [12]

Figure 3.10 shows how the end-to-end delay saturates at a high link transmission rate in telephone transmission.

The end-to-end delay T_a at the transmission rate of 150 Mb/s is substantially determined by the cell assembly/disassembly time and the link propagation delay.

In case of signalling information transfer the delay of the network is important for the connection set-up delay. CCITT Recommendation I.352 gives the provisional values for the connections set-up delay in a 64 kbit/s ISDN connection. ATM switching is very fast, and the delay in one switching element is likely to be less than 1 ms, although it may vary with the traffic load on the switch because of the queuing delays [13]. Assuming that the processing delays in ATM exchanges are comparable to the processing delays in existing digital exchanges, the delay values in Recommendation I.352 are also valid for ATM services, if the same network architecture is considered.

The propagation and processing delays are important parameters for choosing an efficient error correction method in the AAL.

3.4.3 ATM performance evaluation

In this paragraph the ATM performance will be compared with the performance objectives stated in G.821 and shown in Table 3.1.

In the ATM network bit errors and burst errors can cause cell loss. Therefore one have to compare the cell loss at the ATM layer with the bit errors described in G.821. The following assumptions will be made on the ATM layer performance:

- the probability of cell loss caused by random bit errors will be better than 10^{-9} (figure 3.8 and 3.9);
- the burst errors on the physical link occur once in every four days [10]
- during a burst, the BER is close to 0.5 and every cell hit by a burst will be lost.

An errored second with a BER worse than 10^{-3} only occurs at a burst error. Assuming that 2 errored seconds in every four days occur gives a percentage of $4.6 * 10^{-4}$%, which is much better than the objective 0.2% in table 3.1. The same assumption can be made about the degraded minutes. One burst in every four days will cause a maximum of 2 degraded minutes in that period, which results in a percentage of approximately 0.03%. This percentage is much better than the objective of 10% in table 3.1. The third objective in table 3.1, the 92% error free seconds, is easily achieved by the ATM network.

Another cause of cell loss not described in this chapter is buffer overflow due to congestion in the ATM network.

3.5 Conclusions

The AAL for signalling has to provide at least the same service as the DSS1 Data Link Service and the MTP Level 2 Service. The AAL protocol for signalling should be usable at both the UNI and NNI.

The AAL will enhance the service delivered by the ATM and the Physical Layer. The errors on the Physical Layer are predominantly single bit errors with a mean BER less than 10^{-6} and burst errors lasting 20-40 ms and occurring once in every four days.

If the random BER of the physical link is less than 10^{-6}, the expected cell loss rate will be less than 10^{-9}. The ATM performance in terms of cell loss is much better than the performance objectives stated for digital connections of an ISDN.

The dominant factors of the cell delay in the ATM network are the propagation delay and the processing delay. The connection set-up delay values recommended in I.352 will be achieved by the ATM network. The propagation and processing delays are important parameters for choosing an efficient error correction method in the AAL.

Chapter 4

Services and Functions of the AAL

4.1 Introduction

In chapter 3 the starting points for developing an AAL for signalling are described. Another starting point of this chapter will be Report R34 of Study Group XVIII [1], especially the Draft Recommendations concerning the AAL: I.362 and I.363. In paragraph 4.2 and 4.3 the services and functions of the AAL for signalling are described. The set of AAL service primitives will be described in paragraph 4.4.

4.2 Service Description

To pave the way for future work on protocols of higher layers, it is necessary to provide a clear definition on the AAL boundary to the higher layers. This means specification of the AAL-SAP and describing the services provided by the AAL to the layers above. Therefore it is necessary to take into account:

- the protocolstack in the control plane in figure 4.1;
- the positioning of the AAL SAP in the OSI-RM.

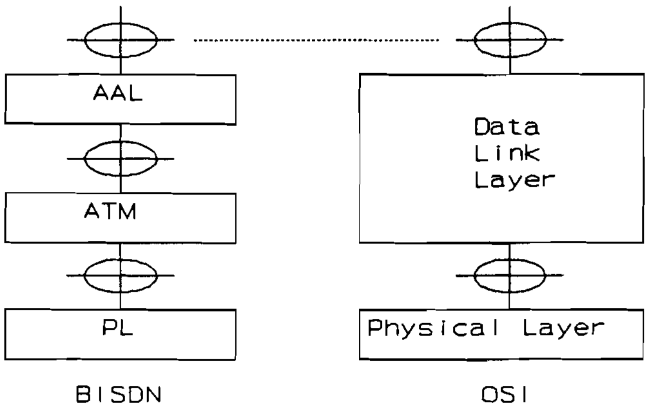


Figure 4.1: Positioning of the AAL SAP in the OSI-RM

In figure 4.1 the positioning of the AAL in the OSI-RM is shown.
As shown in figure 4.1 the AAL has to deliver the same service as the Data Link Service as listed in recommendation X.212 [14]. Therefore the starting point of the AAL service description is the Data Link Service of the OSI-RM.

4.2.1 AAL for signalling Service

The term AAL Service in this paragraph is meant to be the AAL for signalling Service. The components of this AAL Service, that are listed below, are derived from:

- the Data Link Service of the OSI-RM
- the DSS1 Data Link Service
- the MTP level 2 'service'

The AAL Service provides for the transparent and reliable transfer of data between AAL Service users. It makes invisible to these AAL Service Users the way in which supporting communications resources are utilized to achieve this transfer.

In particular, the AAL Service provides for the following:

a) Independence of underlying Layers
The AAL Service relieves AAL Service users from all concerns regarding which configuration is available (e.g., point-to-point connection) or which physical facilities are used.

b) Transparency of transferred information
The AAL Service provides for the transparent transfer of AAL Service user-data. It does not restrict the content, format or coding of information, nor does it ever need to interpret its structure or meaning.

c) Reliable transfer of data
The AAL Service relieves the AAL Service user from loss, insertion, corruption and misordering of data which may occur. In some cases of unrecoverable errors in the ATM Adaptation Layer, duplication or loss of SDUs may occur.

d) Establishment and release of AAL connections
The AAL service makes available to the AAL Service user a means to establish and release AAL connections. Alternatively the AAL service provider may release the AAL connection. Depending on the conditions release of an AAL connection may result in loss of AAL user data.

e) Link Control Service
The AAL service makes available to the AAL service user the means to control an AAL connection.

f) Flow control
The AAL service makes available to the AAL Service user the means by which the receiving AAL Service user may flow control the rate at which the sending AAL Service user may send SDUs.

(In paragraph 4.2.3 a distiction will be made between assured and nonservice, assured service will provide services a) through f) non-assured service will only privide a) and b))

4.2.2 Sublayering of the AAL

Because some interdependent functions must be performed in the AAL, a separation is made into two logical sublayers. These are the Convergence Sublayer (CS) and the Segmentation and Reassembly sublayer (SAR). In figure 4.2 the hierarchy of these layers is shown for message mode operation (paragraph 4.2.3) with an AAL-SDU smaller than the maximum CS-PDU length.

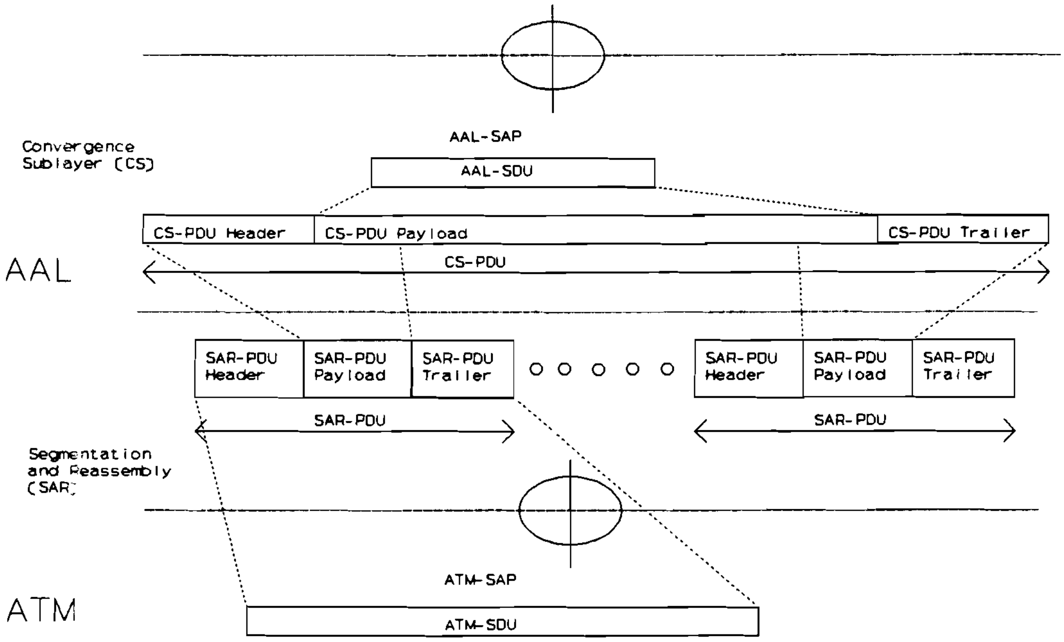


Figure 4.2: Hierarchy and Naming for the AAL

The prime functions of the SAR and CS are:

SAR: - segmentation of higher layer info into a size suitable for the information field of the ATM cell.
- reassembly of the content of ATM cell information field into higher layer information.

CS: - to provide the AAL Service at the AAL-SAP. This sublayer is service dependent.

There are no SAPs defined between these two sublayers. A combination of SAR and CS provides the service needed for signalling.

4.2.3 Service Modes and Delimiting Modes

In Q.921 the service of the data link layer is divided into unacknowledged and acknowledged information transfer service.

In Draft Recommendation I.363 there is also a distinction made between two modes of service:

Assured: Every assured AAL-SDU is delivered with exactly the data content that the user sent. The assured service is provided by retransmission of missing and corrupted CS-PDUs. Flow control is provided as a mandatory feature.

Non-assured: Integral AAL-SDUs may be lost or corrupted. Lost and corrupted AAL-SDUs will not be corrected by retransmission. An optional feature may be provided to allow corrupted AAL-SDUs to be delivered to the user.

Delimiting is an important protocol mechanism. It makes possible for the receiving entity to recognize the boundaries of the transmitted data units. In recommendation I.363 two modes of delimiting are described. These two modes are called 'modes of service'. However, it is better to treat these modes in an OSI-like manner: the (SDU-)delimiting modes are a way in which the AAL protocol supports the AAL service, but they are not an AAL service on itself. The two described SDU-delimiting modes are

- message mode
- streaming mode

Message mode is a form of segmenting/reassembling. An AAL-SDU will be transported in one or more CS-PDUs.
Streaming mode is a form of blocking/deblocking. One or more AAL-SDUs will be transported in one CS-PDU.

For signalling only the message mode will be used. The AAL-SDU will be segmented if the maximum CS-PDU length is reached (depends on protocol considerations and the length of the messages of DSS1 Network Layer and MTP level 3).

4.3 Functions of the AAL

The functions of the AAL can be derived from the functions of Q.921, Q.703 and additional functions required because of the fact that the ATM Layer is used as the transfer mode for B-ISDN.

The following list of functions are necessary for the AAL for signalling:

Segmentation and reassembly
The information field of the ATM cell has a length of 48 octets, but the length of the messages of the higher layers can have a length of 270 octets or more in future higher layer protocols. Therefore a segmentation and reassembly function is necessary to make the signalling messages suitable for the ATM cells.

Service Data Unit delimiting
This function is described in paragraph 4.2.3: the message mode SDU-delimiting. This function can be compared with the use of flags in Q.921 and Q.703.

Error detection
Error detection is necessary to provide reliable transfer of information. It is used for detecting out-of-sequence, loss, inserted and corrupted data.

Error correction
Error correction is necessary to provide the assured mode service.

Link initialization

The link initialization function is appropriate to both first time initialization and initialization after a link failure.

Error monitoring

This function is needed to support the functions in MTP level 3. If the error rate is too high the link will be taken out-of-service and the signalling messages will be redirected.

Flow control

Flow control can be split into two separate functions. These are peer-to-peer and layer-to-layer flow control. Flow control is necessary in the assured service mode.

Link Status Control

This functions have to replace the link status control functions of MTP level 2.

AAL-SDU retrieval

This function is necessary to support the procedures change-back and change-over at MTP level 3. The CS-PDUs at the AAL layer have to be returned to MTP level 3 when a link failure occurs.

These list of functions can be split into functions performed by the SAR or by the CS.

CS functions:

- AAL SDU retrieval
- AAL SDU delimiting
- Error correction
- Error monitoring
- Link initialization
- Flow control
- Link Status Control

SAR functions

- Segmentation and reassembly
- Error detection

Functions of LAPD and MTP level 2 that need not to be performed by the AAL are:

- Addressing/TEI management
- Multiplexing
- Flag synchronization (bit stuffing)

4.4 Service Primitives

The service primitives of the AAL are mainly based upon the service primitives of LAPD. The MTP of SS7 has no OSI-like layered structure. The interaction between MTP level 2 and 3 occurs by messages instead of service primitives. Therefore a SAP, as indicated in figure 4.3, has to be defined.

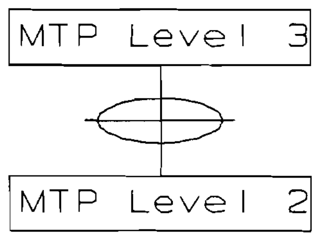


Figure 4.3: Service Access Point between MTP level 2 and 3

In Appendix A the interaction is shown between MTP level 3 and level 2. It is important to define a small number of primitives. The service primitives of LAPD and the messages of the MTP result together in the service primitives of the AAL. In table 4.1 the AAL-primitives derived from the DL-primitives of LAPD are listed.

The messages used for the interaction between level 3 and 2 have to be mapped with the AAL primitives. In order to achieve compatibility

Table 4.1: Primitives derived from Q.921(2)

AAL Primitive	DL primitives	Types			
		Req.	Ind.	Res.	Con.
AAL-ESTABLISH	DL-ESTABLISH	X	X	-	X
AAL-RELEASE	DL-RELEASE	X	X	-	X
AAL-DATA	DL-DATA	X	X	-	-
AAL-UNITDATA	DL-UNITDATA	X	X	-	-
MAAL-ERROR	MDL-ERROR	-	X	-	-

between the AAL and MTP level 3 two additional primitives are defined:

- AAL STATUS
- AAL RETRIEVE

The AAL STATUS primitive is neccesary to support the status information flow between level 2 and 3. The AAL RETRIEVE primitive is necessary in case of a link failure, because then all unacknowledged messages are retransmitted by MTP level 3 on another signalling link.

With these two primitives it is possible to map all the MTP messages in Appendix A with types of AAL primitives. (table 4.2)

The types of the AAL primitives are listed in table 4.3.

With this table of primitives a uniform (UNI and NNI) state transition diagram has to be developed for AAL connection endpoints.

Table 4.2: Mapping of MTP messages with AAL primitives

MTP message	AAL primitive	type
Start	ESTABLISH	req.
Stop	RELEASE	req.
In-Service	STATUS	ind.
Out-of-Service	STATUS	ind.
Local processor outage	RELEASE	req.
Remote processor outage	RELEASE	ind.
Local processor recovery	RELEASE	req.
Remote processor recovery	ESTABLISH	ind.
Transmit data	DATA	req.
Receive data	DATA	ind.
Retrieve BSNT	RETRIEVE	req.
Retrieval request and FSNC	RETRIEVE	req.
Retrieved message	RETRIEVE	ind.
Retrieval complete	RETRIEVE	ind.
Retrieved message BSNT	RETRIEVE	ind.
Congestion	STATUS	ind.
Congestion ceased	STATUS	ind.

Table 4.3: AAL primitives and types

AAL Primitive	Types			
	Req.	Ind.	Res.	Con.
AAL-ESTABLISH	X	X	-	X
AAL-RELEASE	X	X	-	X
AAL-DATA	X	X	-	-
AAL-UNITDATA	X	X	-	-
AAL-STATUS	X	X	-	-
AAL-RETRIEVE	X	X	-	-
MAAL-ERROR	-	X	-	-

4.5 Conclusions

The AAL service is a combination of the service provided by LAPD and MTP level 2. A starting point for this service description was also the Data Link Service of the OSI-RM. For supporting this service the AAL is separated into two sublayers, the Convergence Sublayer and the Segmentation and Reassembly sublayer.

The functions of the AAL are separated in functions of the CS and functions of the SAR. SAR functions are segmentation and reassembly and error detection. Some CS functions are error correction/monitoring, flow control, link status control.

A uniform set of primitives for both UNI and NNI is proposed. A SAP is defined between MTP level 2 and 3 by mapping the MTP messages to the AAL service primitives. Two primitives, AAL-RETRIEVE and AAL-STATUS are defined to perform the functionality needed by the signalling network management at MTP level 3.

Chapter 5

SAR and CS

5.1 Introduction

In I.362 [1] a service classification for the AAL is made to minimize the number of AAL protocols. This service classification (A..D) is based on the following parameters:

- timing relation between source and destination;
- bit rate;
- connection mode;

In I.363 [1] different AAL types are defined to support the defined service classes in I.362. An AAL type specification is an AAL protocol which consist of a combination of SAR and CS functions. For signalling only AAL type 3 or type 4 could be used. In this paragraph the coding of the SAR-PDU and the CS-PDU of AAL type 3 will be discussed. AAL type 3 supports the service class which requires no timing relation between source and destination, has a variable bit rate and is connection-oriented (Class C).

5.2 Structure and Coding of SAR-PDU

5.2.1 SAR in R34

Some fields of the SAR-PDU are defined in R34 [1]. The error detection function and the segmentation and reassembly function of the SAR sublayer requires a:

- Segment type (ST)
- Sequence number (SN)
- Reserved (RES)
- Length indicator (LI)
- Cyclic redundancy check (CRC)

In figure 5.1 the coding of the SAR-PDU as defined in I.363 is shown:

ST	SN	RES	SAR-PDU Payload	LI	CRC
2	4	10	44 octets	6	10

Figure 5.1: SAR-PDU coding defined in I.363 [1]

The Segment Type field is used to indicate the location of the SAR-PDU payload in the CS-PDU (message). The ST field can indicate, in the case of CS-PDUs longer than 44 octets, Begin Of Message (BOM), Continue Of Message (COM), End Of Message (EOM). When the CS-PDU is smaller than 44 octets the ST field will indicate SSM (Single Segment Message).

The REServed field is reserved for further use.

The LI field is encoded with the number of octets of the CS-PDU that are included in the SAR-PDU payload field.

The CRC field is used to detect errors in the SAR-PDU including the SAR-PDU header, SAR-PDU payload and the LI field. For the calculation of this field a CRC 10 generating polynomial $G(x)=1+x+x^4+x^5+x^9+x^{10}$ is used. This CRC polynomial is capable to detect:

- all single bit errors
- all double bit errors
- all odd numbers of errors
- all burst errors of length 10 bits or less
- a fraction $1 - 1/2^9 \approx 0.99805$ of the burst errors of length 11
- a fraction $1 - 1/2^{10} \approx 0.99902$ of the burst errors of length greater than 11

A SAR-PDU with 4 bit errors (not a burst) will not be detected. The probability of such SAR-PDU due to random bit errors is very low. In figure 3.6b is shown that this probability is approximately 10^{-15} at a random BER of 10^{-6} and 10^{-19} at a BER of 10^{-7}. A probability of 10^{-15} at a 150 Mb/s link means that one errored SAR-PDU in 89 year will be undetected! So, the number of errored SAR-PDUs that remain undetected after the CRC will be very small.

The number of bits of the sequence number is important for the error detecting capabilities. When the SN has a length of 4 bits the SAR sublayer can detect errors in CS-PDUs with a maximum length of $2^4 \times 44 = 704$ octets.

Due to cell loss it is possible that mis-assembling of a CS-PDU occur. In figure 5.2 the effect of the loss of several SAR-PDUs is shown.

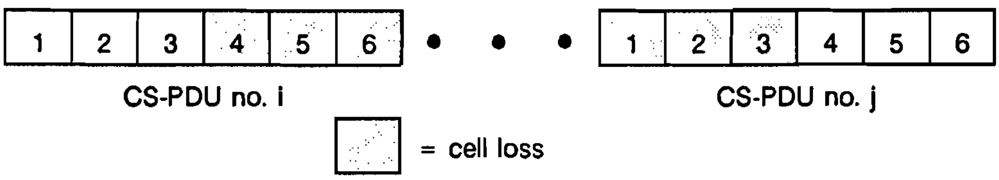


Figure 5.2: Mis-assembling of a CS-PDU due loss of several SAR-PDUs

Mis-assembling as shown in figure 5.2 occurs when the SN of a correct received SAR-PDU, belonging to a certain CS-PDU i matches with the SN of the next correct SAR-PDU, belonging to a CS-PDU j.

The probability of this mis-assembling will be very low, but it can happen for instance during a burst error. A solution for this can be a larger sequence number than four bits. With a sequence number of for example 14 bits is it possible to detect a transmission interrupt of more than 40 ms (150 Mb/s).

5.2.2 Additional SN in SAR-PDU

In order to avoid the misassembling shown in figure 5.2 the reserved field in figure 5.1 can be used. The CS-PDUs will be numbered with an 8 bit sequence number. In figure 5.3 the reserved field is used for:

- a larger SN (6 bits)
- a SN of the CS-PDU (CSSN)

to enhance the detecting capabilities of the SAR sublayer.

ST	SN	CSSN	SAR-PDU Payload	LI	CRC
2	6	8	44 octets	6	10

Figure 5.3: SAR-PDU structure

It is expected that message length for B-ISDN call will become longer than for N-ISDN call to handle complicated connection configurations such as multi-point or multi-connection. Therefore, it is preferable that the maximum CS-PDU informatiom field size is at least 1k octets [15]. The sequence number in figure 5.3 is 6 bits long. So, the SAR sublayer can handle CS-PDUs with a maximum length of $2^6 * 44 = 2816$ octets. This length will be long enough for signalling messages of todays and for future higher layer signalling protocols [15]. The CSSN field is the sequence number of the CS-PDU. With this sequence number in every SAR-PDU the misassembly of figure 5.2 can never occur, if the number of outstanding CS-PDUs is limited. With the sequence number of the CS-PDU 8 bits long the AAL has twice the capacity of MTP level 2 and LAPD.

5.3 SAR state and flow diagram

In this section the actions of the SAR transmitter and receiver will be described with state and flow digrams. For both the transmitter and the receiver two states are defined (table 5.1)

Table 5.1: States of SAR Transmitter and Receiver

	Name State	Description
Transmitter	ST0	awaiting a next CS-PDU
	ST1	segmenting a CS-PDU
Receiver	SR0	awaiting a next CS-PDU
	SR1	assembling a CS-PDU

5.3.1 SAR Transmitter

An easy representation of the states and the transitions in the transmitter is shown in figure 5.4.

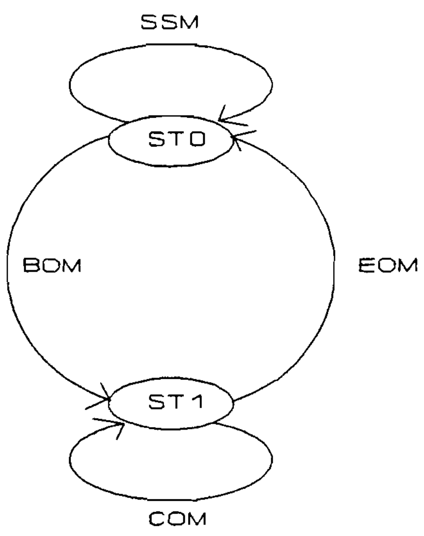


Figure 5.4: State Transition Diagram for a SAR Transmitter

For a clear definition of the actions that take place in the transmitter several variables are defined.

Table 5.2: Variables of SAR Transmitter

S(SN)	Sequence number of the transmitted SAR-PDU
S(CSSN)	CS sequence number of the transmitted SAR-PDU
S(LI)	LI field of the transmitted SAR-PDU
VS(SN)	Next sequence number
VS(CSSN)	Next CS sequence number
VS(LI)	Remain of the transmitting CS-PDU
L	Length of the CS to be transmitted

The actions of the transmitter are less complex if the length of the transmitted CS-PDU is smaller that 44. This is shown in figure 5.5 where the specific actions are shown by an SDL diagram.

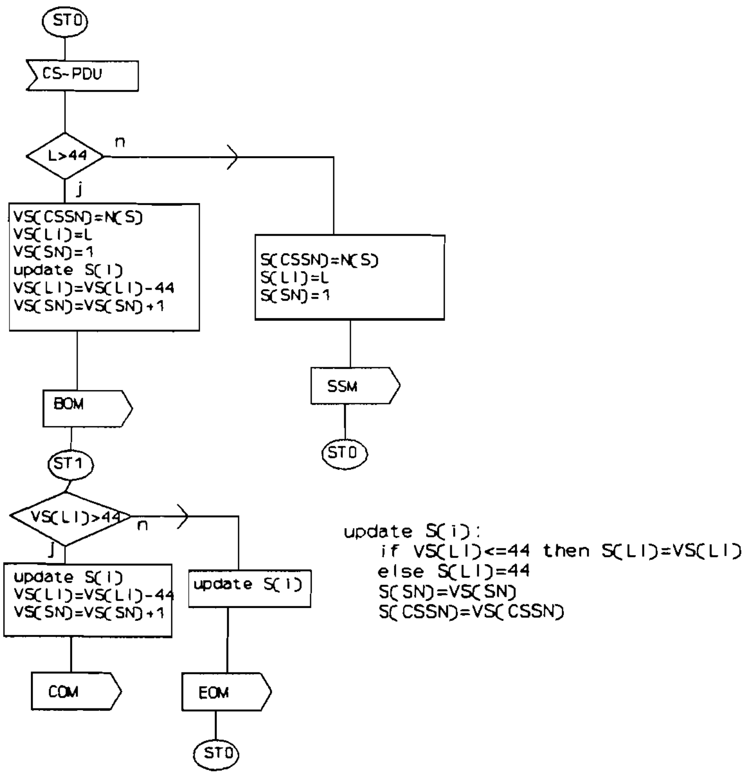


Figure 5.5: SDL diagram of SAR Transmitter

5.3.2 SAR Receiver

The state transition diagram of the receiver, as shown in figure 5.6, is more complex than the STD of the transmitter.

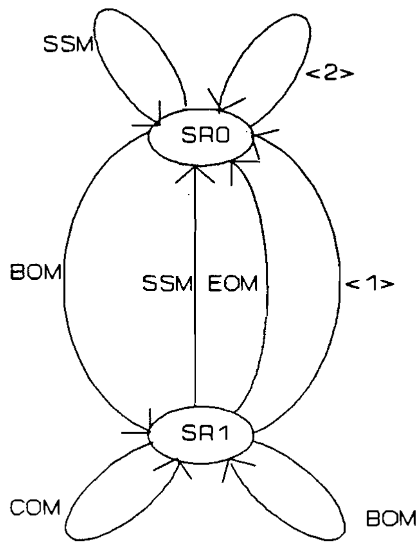


Figure 5.6: State Transition Diagram for a SAR Receiver

The transitions in figure 5.6 marked with an ST indicate a <u>correct</u> received SAR-PDU. In state SR0 the receiver waits for a correct SSM or BOM SAR-PDU. The transition marked with <2> means that an incorrect, EOM or COM SAR-PDU is received. In that case the receiver discards the SAR-PDU and returns into state SR0.

In normal operation, no cell loss and insertion at the ATM layer and no bit errors in the SAR-PDU, the receiver in SR1 waits for a COM or EOM SAR-PDU to assemble the CS-PDU. If a SSM or a BOM SAR-PDU is received in state SR1, all the received SAR-PDUs of the CS-PDU that is assembled are discarded. So, the receiver does not take specific actions to overcome inserted ATM cells. The reason for this is the complexity and the low probability of inserted cells. The transition marked with <1> means that an incorrect SAR-PDU is received.

The defined variables for the SAR receiver are shown in table 5.3. With these variables the SDL diagrams shows the specific actions of the receiver.

Table 5.3: Variables of SAR Receiver

R(SN)	Sequence number of the received SAR-PDU
R(CSSN)	CS sequence number of the received SAR-PDU
R(LI)	LI field of the received SAR-PDU
VR(SN)	Next sequence number
VR(CSSN)	Next CS sequence number

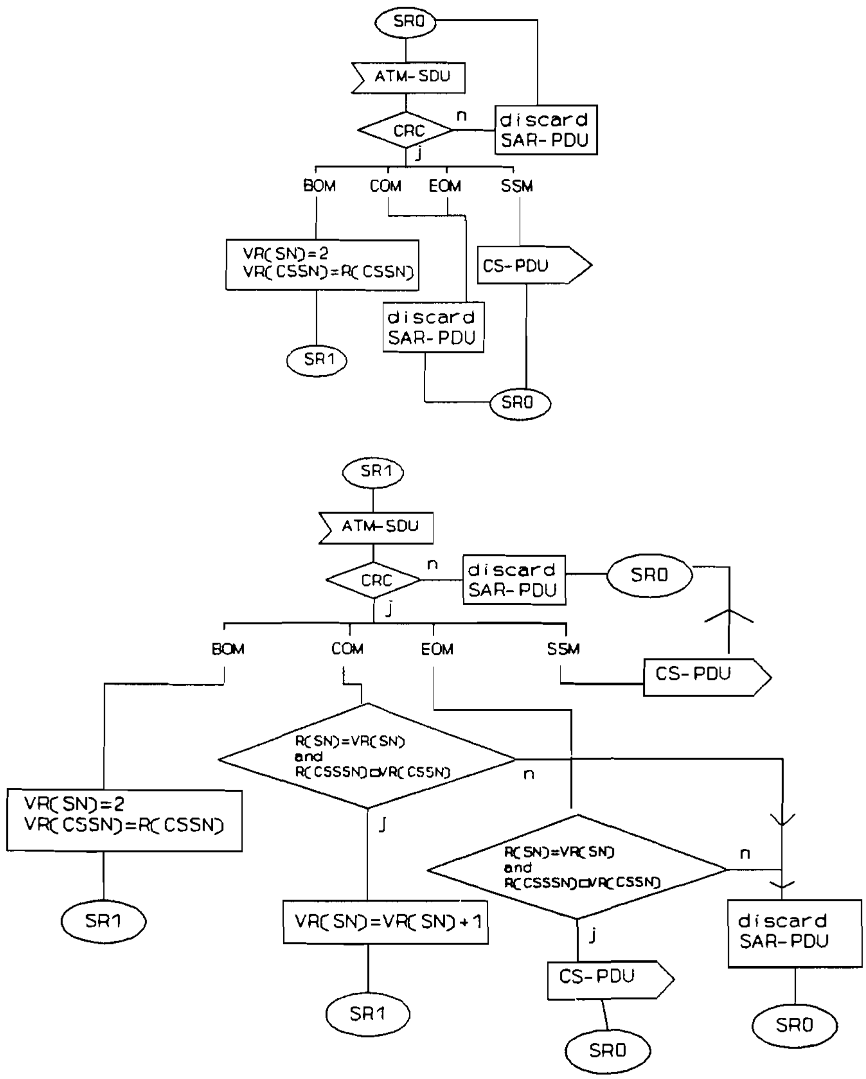
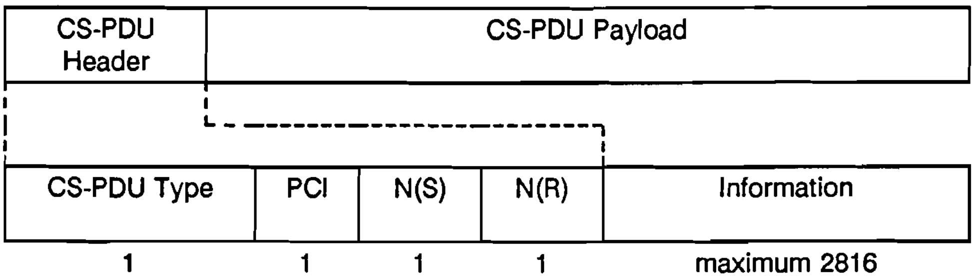


Figure 5.7: SDL diagram of SAR Receiver

5.4 Structure and Coding of CS-PDU

The CS provides a signalling link between two signalling entities and must be able to provide the same layer service as provided by LAPD and MTP level 2. In paragraph 3.2.1 is proposed that the same protocol should be applicable at both the UNI and NNI. Therefore it is necessary to choose a symmetrical AAL protocol. The data units described in this paragraph are based on the Q.922 protocol (HDLC-based) in order to ease the mapping of (and interworking with) HDLC-based protocols which are currently widely used for signalling information transfer. Figure 5.8 shows the general coding of the CS-PDU. (general information about HDLC is given in Appendix B)



PCI = Protocol Control Information (C/R, P/F etc.)
N(S) and N(R) have comparable functionality as in HDLC

Figure 5.8: Convergence Sublayer Coding

5.4.1 Different data units

The control field of the data units of a HDLC-based protocol has three different formats

- Information (I)
- Supervisory (S)
- Unnumbered (U)

These different formats will also be used by the CS-protocol. Data units of MTP level 2 have also three different formats: MSU, LSSU and FISU.

The MTP functions will be replaced by the HDLC-based AAL-protocol functions. Therefore the data units will be compared with the formats of a HDLC protocol.

MSU The function of the MSU in MTP level 2 is the same as the function of the I frame in HDLC.

FISU The FISU is used for the error monitoring. This FISU can be replaced by a signal unit with a supervisory format control field.

LSSU The LSSU are used to inform the remote signalling point of the status. These functions can be fulfilled by signal units of supervisory or unnumbered format.

In table 5.4 the mapping between the MTP signal units and the HDLC signal units is shown.

Table 5.4: Mapping between MTP level SU and HDLC frames

MTP level 2 SU	HDLC format	type
MSU	I	
FISU	S	EM (table 5.5)
LSSU:		
O: out of alignment	not used	-
N: normal alignment	not used	-
E: emergency alignment	not used	-
OS: out of service	U	DM
PO: processor outage	U	DISC
B: busy	S	RNR

Figure 5.9 shows the three different formats of the CS-PDU header.

I	CS-PDU Type	0 0 0 0 0 0 0 P	N(R)	N(S)
S	CS-PDU Type	S S S S S S C P R F	N(R)	-
U	CS-PDU Type	M M M M M M C P R F	-	-

N(S)	Transmitter send sequence number
N(R)	Transmitter receive sequence number
P/F	Poll/Final bit
C/R	Command/Response bit
S	Supervisory fuction bit
M	Modifier function bit

Figure 5.9: I, S and U format of the CS-PDU

The S bits in figure 5.9 identify the supervisory PDU formats and are listed in table 5.5.

Table 5.5: Coding of the supervisory CS-PDU format

SSSSSS	Abbr.	Name
000001	RR	Receive Ready
000010	RNR	Receive Not Ready
000100	SREJ	Selective Reject
000101	EM	Error Monitoring

In relation to the supervisory formats of LAPD two extra formats are listed in table 5.5:

- Selective Reject
- Error Monitoring

The SREJ format is used for to perform the error correction function. The function will be explained in chapter 6. The EM format is used to perform the error monitoring neccessary for MTP level 3.

The M bits in figure 5.9 identify unnumbered PDU formats and are listed in table 5.6.

Table 5.6: Coding of the unnumbered CS-PDU format

MMMMMM	Abbr.	Name	C/R
000110	SABMM	Set Asynchronous Balanced Mode Mod256	C
000111	DM	Disconnect Mode	R
001000	UI	Unnumbered Information	C
001001	DISC	Disconnect	C
001010	UA	Unnumbered Acknowlegement	R
001011	FRMR	Frame Reject	C
001100	XID	Exchage Identification	C/R

The SABMM is a command to place the peer entity into modulo 256 multiple data unit operation.

5.5 Conclusions

AAL type 3 has to be used for signalling. The structure and coding of the SAR-PDU and the CS-PDU are described in this chapter. It is show that almost al errors in the SAR-PDU are detected by the CRC. The SN of the SAR-PDU is expanded with 2 bits to 6 bits. The maximum protected CS-PDU length is therefore 2816 octets. This length is appropriate to transport all signalling messages [15] (at least 1k). A CSSN field in the SAR-PDU is ensures that no misassembling of CS-PDUs will occur.

The coding and structure of the CS-PDU is based on the HDLC protocol. The signal units of the MTP level 2 are mapped with the I, S, and U formats of the CS-PDU. A SABMM command is defined which is used to set the peer entity in multiple data operation with 8 bit sequence numbering. EM is an extra supervisory CS-PDU to perform the error monitoring when no I-CS-PDUs are available or the connection is released.

Chapter 6

AAL Protocol and Procedures

6.1 Introduction

In chapter 4 an uniform set of primitives is defined for the AAL. Based on this set of primitives a uniform state transition diagram for both the UNI and NNI will be shown in paragraph 6.2.

The functions error correction and error monitoring of the AAL will be explained in paragraph 6.3 and 6.4. The flow control function is described is paragraph 6.5.

6.2 State Transition Diagram

In figure 6.1 a uniform state transition diagram (STD) for both the UNI and NNI, is shown. Because the AAL protocol is derived from LAPD, the starting point of the STD in figure is the STD of Q.921.

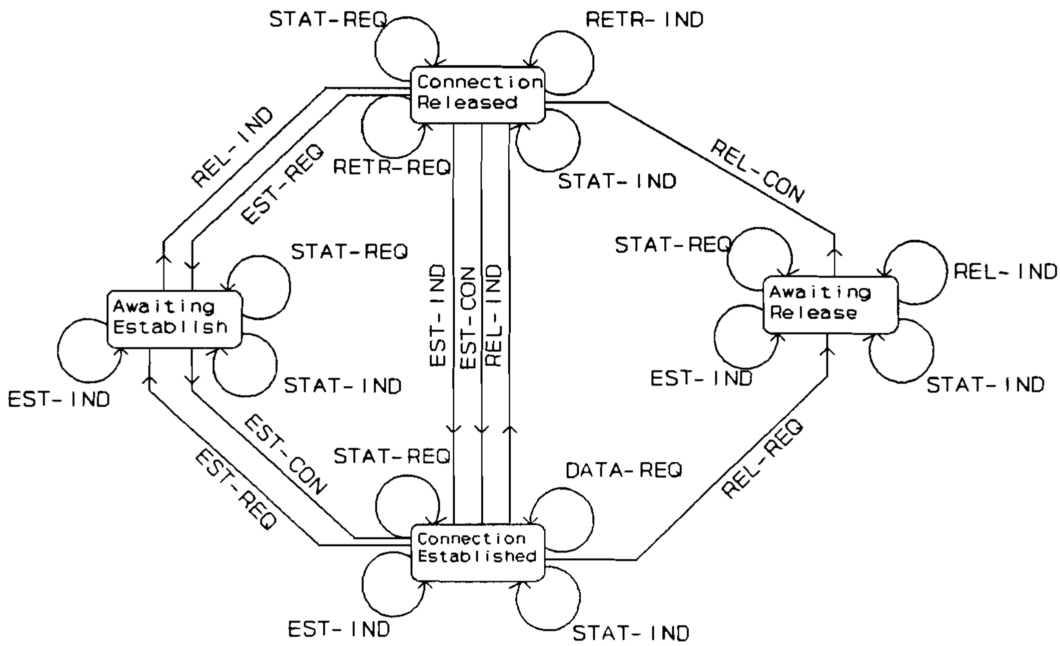


Figure 6.1: State Transition Diagram for sequences of primitives at a point-to-point AAL Connection as seen by Layer 3

In MTP level 2 two states are defined that are not necessary in the AAL:
- Await alignment state
- Processor outage state

The function 'initial alignment' is replaced by the establishment procedure. During this procedure no 'proving period' is necessary.

If a local or remote processor outage condition occurs the link will be released. (This is also shown by the mapping of the primitives in table 4.2: the local processor outage message is mapped with a RELEASE-REQUEST.)

The primitive STATUS-INDICATION can be send to layer 3 in every state. Optional MTP level 3 can ask for STATUS-INDICATION by a STATUS-REQUEST.

The RETRIEVE-REQUEST primitive is sent from layer 3 when the AAL connection is released due to a link failure or a remote processor outage condition.

6.2.1 Meta-signalling

Before the sequences of primitives will be described, something has to be explained about meta-signalling. In Q.921 TEI management provides the addressing of the signalling endpoints. In B-ISDN a meta-signalling protocol will be developed to provide a comparable service.

In SS7 the transport of signalling messages takes place over a signalling network. It is agreed in a recent CCITT SG XI meeting that the NNI signalling in B-ISDN will be based on point-to-point ATM connections. The signalling virtual channel identifier (SVCI) in this case will
be standardized with VCI=0, VPI=2 [17]. So there is no need for addressing multiple endpoints at the UNI.

At the UNI a meta-signalling protocol will be defined that manages the point-to-point signalling virtual channels. The meta-signalling protocol provides the assignment and removal of point-to-point signalling virtual channels to a signalling endpoint and checks the status of the signalling channels. The meta-signalling uses a virtual channel, which is identified by a standardized value of VPI=0 and VCI=1. Meta-signalling is located in the layer management of the ATM layer. [17]

A general broadcast signalling virtual channel is identified by a standardized value of VPI=0, VCI=2.

6.2.2 Sequences of Primitives

In this paragraph several sequences of primitives will be described. In figure 6.2 the sequence of primitives in case of successful connection establishment is shown.

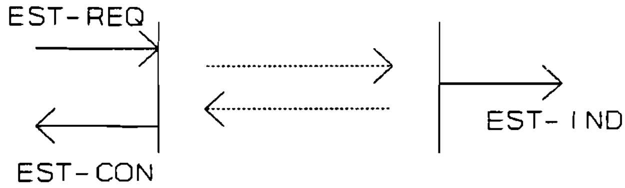


Figure 6.2: Sequence of primitives in case of successful connection establishment

In case of unsuccessful connection establishment, due to local or remote causes the ESTABLISH.CON primitive will be replaced by a RELEASE.IND with a parameter that informs layer 3 about the reason for unsuccessful connection establishment.

The sequence of primitives in case of connection release are also shown in figure 6.2 with ESTABLISH replaced by RELEASE.

In figure 6.3 the sequences of primitives are shown in case of data transfer, retrieve and status procedure.

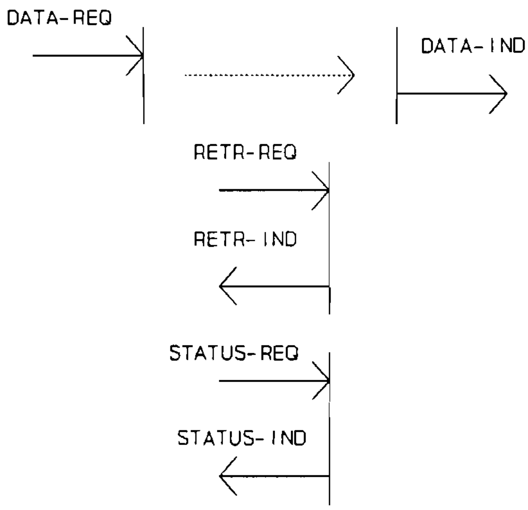


Figure 6.3: Sequences of primitives in case of data transfer, retrieve and status procedure

The RETRIEVE-REQ primitive has no parameter. As shown in Appendix A and table 4.2 MTP level 3 will ask at the beginning of a buffer updating procedure for the:

- BSNT	The sequence number of the last message that has been correctly received.
- MESSAGE+FSNC	The messages that have to be retransmitted at another signalling connection and the sequence number of the next message to be transmitted.

These two messages will be mapped with one RETRIEVE-REQ primitive and the response of the AAL will be several RETRIEVE-IND primitives with all the necessary information.

The RETRIEVE-IND primitive has a parameter *type*. This *type* parameter indicates in the following order:

- BSNT
- FSNC
- message
- complete

First the sequence numbers will be sent to level 3, then the messages and an indication that all messages are send. The abbreviations BSNT and FSNC are MTP level 2 abbreviations. In the AAL protocol they will be replaced by other names of variables.

The STATUS-IND primitive in figure 6.3 has two parameters: *state* and *cause*. In Q.704 is described that a signalling link is always considered by MTP level 3 in one of two possible states. The parameter *state* has therefore two possible codes: available or unavailable. If the parameter *state* indicates that the AAL connection is unavailable the parameter *cause* can have the following codes:

- failed or inactive
- blocked
- (failed or inactive) and blocked

An AAL connection is recognized as failed if a connection failure condition occurs (see paragraph 6.4.1). An AAL connection is blocked when a remote processor outage condition exist.

6.3 Error Correction

In paragraph 2.4.3 is shown that the cell loss due to random bit errors will be in the order of at least 10^{-9}. In paragraph 3.3.1 the performance parameter for the loss of messages guaranteed by the MTP is described (10^{-7}). Therefore one could conclude that no error correction function is necessary in the AAL.

However, at the ATM layer burst errors occur due to protection switching and due to congestion (buffer overflow). To overcome the loss due to these burst errors, error correction is a necessary function in the AAL.

(The error detection function is completely located in the SAR layer and is performed by CRC. If the outcome of the CRC is negative the SAR-PDU will be discarded.)

6.3.1 Error correction methods

In general there are two different methods for error correction [18]: Forward Error Correction (FEC) and Automatic Repeat Request (ARQ). In the case of FEC the receiver attempts to correct errors by the use of correction codes (send as overhead in the message). An ARQ simply detects the errors and the transmitter retransmits erroneous messages.

ARQ requires extra overhead for retransmission requests and retransmitting messages in error, but total overhead is often far less than it would be with FEC. Several variants on the ARQ are common [18]:

- Stop-and-Wait
- positive and/or negative acknowledgements
- go-back-N
- selective reject
- piggy-backed acknowledgement

The error correction methods that are used by the ISDN signalling protocols at layer 2 will be classified below.

As shown in paragraph 2.2.1, the Basic Method and the Preventive Cyclic Retransmission method are used by MTP level 2 for error correction. Q.921 uses one method of error correction. In table 6.1 the error correction methods of ISDN signalling protocols are classified.

Table 6.1: Error correction methods in ISDN

BM (MTP level 2)	go-back-N positive/negative acknowledgement piggy-backed acknowledgement
PCR (MTP level 2)	positive acknowledgement piggy-backed acknowledgement retransmission of unacknowledged messages
Q.921	go-back-N positive/negative acknowledgement piggy-backed acknowledgement

So, the signalling protocols in ISDN at both the UNI and NNI uses the go-back-N method with positive and negative (piggy-backed) acknowledgement. The use of the negative acknowledgements is not necessary, but they can speed retransmissions since they may be returned before the original transmitter would finish timing out.

6.3.2 Satellite links

Satellite links will also be integrated in the B-ISDN. This integration has consequences for the retransmission method of the error correction function.

In [20] a motivation is given for the use of selective retransmission. At high link speeds the use of the go-back-N method causes a degradation in both throughput and delay. Furthermore, if the loss of a message is due to mild congestion, the spurious retransmissions will aggravate the congestion giving rise to still more messages being lost causing more spurious retransmissions and so on. Figure 6.4 shows the degradation in throughput for the go-back-N protocol operating over a 45 Mb/s satellite connection.

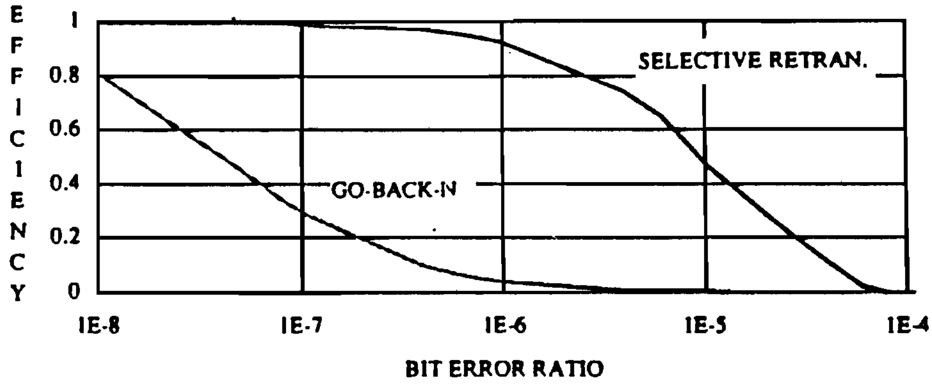


Figure 6.4: Efficiency of a 45 Mb/s Satellite Link for a frame size of 4096 octets [20]

6.3.3 AAL error correction method

In paragraph 6.3.1 the methods of error correction in ISDN signalling protocols at layer 2 are shown. In paragraph 6.3.2 the need of selective retransmission in the case of satellite links is motivated.

In paragraph 4.3 the functions of the AAL are listed and divided into functions performed by the SAR or the CS. There the decision to perform the error correction function at the CS layer was already made. So, the AAL will only retransmit errored CS-PDUs and no errored SAR-PDUs because otherwise the AAL error correction function is too complex.

In paragraph 5.4 in table 5.5 already is shown that the error correction will be based on selective retransmission by use of a SREJ supervisory CS-PDU.

The protocol uses positive and negative acknowledgements. The I CS-PDU can be acknowledged in four different ways:

positive: a supervisory RR CS-PDU, used when no I CS-PDU has to be transmitted

positive: a I CS-PDU (embedded acknowledgement)

negative: a supervisory SREJ CS-PDU to obtain retransmission of the I CS-PDU indicated by this SREJ

The retransmission policy of the AAL is not the same as the policy in LAPD caused by the SREJ option. In paragraph 6.5 the flow control mechanism will be described based on 'sliding windows'. The influence of the SREJ option on the size of the window will be explained.

6.4 Error Monitoring

The AAL signalling connection must have a high reliability. Therefore an error monitoring function is needed in the AAL for the signalling network management at MTP level 3. The purpose of the signalling network management functions is to provide reconfiguration of the signalling network in case of failures by changing the routing of signalling traffic.

Before the error monitoring method in the AAL will be described the link failure conditions of MTP level 2 will be investigated.

6.4.1 Link Failure of MTP Signalling Link

In Q.704 several causes are described which result in an indication from MTP level 2 to level 3 that the signalling link has failed. Some of these causes will also occur at the AAL. The causes can be divided into two subjects.

a) intolerable high signal unit error rate
b) causes which will be recognized by the AAL protocol

Some causes of b) are:

- excessive delay of acknowledgements
- errored sequence numbers
- excessive periods of level 2 congestion

'Cause a)' will be detected by the error monitoring function.

Two basic parameters which determine a link failure situation can be used for the error monitoring:

E_p: lowest allowable long term signal unit error rate (percentage)
N_e: maximum of consecutive errored signal units

6.4.2 Monitoring Method in the AAL

In 6.4.1 parameters are defined for the error monitoring. These parameters can be monitored in the AAL at two 'levels'. E_p and N_e can be monitored at the CS or at the SAR sublayer. An errored SAR-PDU will automatically mean that a loss of a CS-PDU. On the other hand the loss of several SAR-PDUs will not automatically lead to the loss of several CS-PDUs.

Therefore for both the CS and the SAR the parameters E_p and N_e have to be specified. The parameters for the CS can be derived from the MTP level specifications:

$E_p = 4\ \%$
$N_e = 64$

At both the CS and the SAR the layer management will be informed by a MAAL-ERROR-INDICATION primitive. This primitive has two parameters: *layer* and *error*. The combinations of parameters codes are listed in tabel 6.2.

Table 6.2: Combination of parameter codes of MAAL-ERROR-INDICATION

layer	*error*
SAR	CRC error unexpected ST unexpected SN
CS	unexpected N(S)

When the AAL connection has been released due to a link failure the CS will transmit supervisory CS-PDUs of the Error Monitoring format. Then the EM CS-PDUs perform the same function as the FISU of MTP level 2.

6.5 Flow Control

Flow control can be divided into *peer-to-peer flow control* and *interface flow control.* Interface flow control is the 'backpressure' of the service user or service provider after a DATA-REQUEST or DATA-INDICATION. Backpressure can be described as the service equivalent of peer-to-peer flow control. In relation to the AAL only the peer-to-peer flow control will be described.

6.5.1 Sliding Window Flow Control

One of the most common approaches to flow control is sliding window flow control. In this type of flow control, a transmitter is given a permit to transmit a 'window' of data and is not allowed to transmit more until it receives another permit. The receiver can regulate flow control by withholding or granting permits.

This type of flow control is also been used by LAPD and will be used by the AAL protocol.

The maximum size of the window is called W. There is a relation between the throughput of the connection, the propagation delay τ and the maximum size of the window [19]. τ is the one way propagation time in seconds. If the transmitter wants to transmit R CS-PDUs per second, the minimum size of the window have to be $2R\tau$. ($W \geq 2R\tau$)

This relation can be verified at an AAL signalling connection. In chapter 5 is shown that the sequence number of the I-CS-PDU is 8 bits long. With this size of the sequence number, the window will be big enough to reach an efficient throughput.

N(S) will be numbered mod M (M=256). The window, all sequence numbers of transmitted I-CS-PDUs not yet acknowledged, has the size W. For a correct flow control it is necessary that $W \leq M - 1$, because otherwise in the case of a rotation of the window over more than M sequence numbers the transmitter can not interpret the acknowledgments correct.

In [19] is showed that the restriction $W \leq M - 1$ is not enough in the situation that a SREJ option is used. Sequence errors will occur when the transmitter do not send a status request in the case of a time-out, but spontaneous retransmits the unacknowledged I-CS-PDUs (timer is started when a I-CS-PDU is transmitted).

So, to avoid this problem the CS-transmitter have to ask for a status request after time-out.

If only the REJ option would be used by the CS, the receiver only accepts the CS-PDUs in correct sequence. If the SREJ is used, the receiver will not only accept the in-sequence CS-PDU but also the following CS-PDUs.

6.5.2 Congestion

The receiving end of a signalling link in MTP level 2 which detected congestion return a LSSU containing a status indication "B" at an interval of about 80-120 ms.

This function can be replaced in the AAL, as indicated in table 5.4, by a S-CS-PDU Receive Not ready (RNR). After the reception of such a RNR a timer will be started and MTP level 3 will be informed by a AAL-STATUS-INDICATION primitive about the congestion. If the timer expires the AAL connection will be released and MTP level 3 have to take appropriate actions.

6.6 AAL Performance Evaluation

In chapter 3 the overall delay of signalling messages in the MTP was described. This message delay will be lower in the AAL (B-ISDN) due to:

- fiber optic links
- selective retransmission
- higher bit rate

Below the performance of the AAL will be compared with the performance objectives of the MTP (paragraph 3.3.1).

undetected errors
Undetected error in the AAL occur when an errored SAR-PDU is not detected by the CRC check. In paragraph 5.2.1 was shown that, with a random BER of 10^{-6} at the Physical Layer (pessimistic), the probability of an undetected errored SAR-PDUS is about 10^{-15}. This is much better than the performance objective of MTP (10^{-10}).

loss of signalling messages
The probability of retransmission of CS-PDUs will be very low, cause by the low CLR ($< 10^{-9}$) at the ATM layer. The loss of CS-PDUs will only occur if the error correction procedure (retransmission) fails. It is expected that the loss after the error correction function is much better than 10^{-7}.

During the standardization of future higher layer signalling protocols effort have to be made to ensure that a lot of signalling messages can be transported by one ATM cell. Therefore it is necessary that the overhead in the AAL have to be minimized. In the AAL proposed in this report the maximum overhead percentage is 8/48 = 16,6%. (4 octets overhead at both the CS and SAR sublayer). The maximum length of the signalling message that can be transported by one ATM cell is 40 octets.

6.7 Conclusions

An uniform state transition diagram is defined for an AAL connection for both the UNI and NNI. This STD is mainly based on the STD of LAPD. The initial alignment procedure of MTP level 2 will be replaced by the connection establishment procedure. In case of a local or remote processor outage condition the AAL connection will be released.

The error correction function is based on the selective retransmission method (timer expiry and SREJ). The selective retransmission method will enhance the performance of the AAL in relation to LAPD.

An AAL connection have to be monitored at both the CS and the SAR sublayer. The monitoring function will be performed by the AAL management.

Flow control is performed by a window mechanism based on 8 bits numbering of the I-CS-PDUs. In case of congestion the peer entity will be informed by a RNR.

The performance of the AAL (delay, loss, undetected errors) will be much better than the performance objectives of the MTP.

Conclusions

* The ISDN signalling protocols are converging toward a single control protocol. The AAL for signalling replaces the functionality of both LAPD and MTP level 2 and will be common for both the UNI and NNI.

* Cell loss rate of the ATM layer and random bit errors at the Physical Layer are important parameters that influence the AAL protocol. Causes of cell loss are random bit errors, burst errors and congestion.

* The AAL is divided into two sublayers: the Convergence Sublayer (CS) and the Segmentation and Reassembling sublayer (SAR). The proposed coding of the CS-PDU is based on the HDLC protocol. A uniform state transition diagram based on LAPD is defined for both the UNI and NNI.

* The performance of the error correction function is enhanced by the use of selective retransmission. Error monitoring is provided at both the CS and SAR sublayer and will be performed by the AAL management. The performance of the AAL (delay, loss, undetected errors) will be much better than the performance objectives of the MTP.

* During the standardization of future higher layer signalling protocols effort have to be made to ensure that most frequently used signalling messages can be transported by one ATM cell.

References

[1] *General B-ISDN aspects*, Report R.34 of SG XVIII, Geneva meeting 23-25 may 1990

[2] *New directions in signalling for broadband ISDN*
Steven E. Minzer, Dan R. Spears, feb. 1989, IEEE Comm. Mag., pg. 6-14

[3] *Signalling System No.7: A tutorial*
Abdi R. Modarressi, Ronald A. Skoog, july 1990, IEEE Comm. Mag., pg. 19-35

[4] CCITT Blue Book - 1988, vol VI, Fascicles VI.10 and VI.11 *Digital Subscriber Signalling System No.1*, Recommendations Q.920-Q.921 and Q.930-Q.940

[5] CCITT Blue Book - 1988, vol. VI, Fascicles VI.7, VI.8 and VI.9 *Signalling System No.7*, Recommendations Q.700-Q.716, Q.721-Q.766 and Q.771-Q.795

[6] *The Role of ISDN Signalling in Global Networks*
Timothy F. Kearns, Maureen C. Mellon, july 1990, IEEE Comm. Mag., pg. 36-43

[7] CCITT Recommendation: *Message transfer part signalling performance*, Vol. VI, Rec. 706

[8] CCITT Recommendation: *Error performance of an international digital connection forming part of an ISDN*, Vol, III, Rec. G.821

[9] *Observations of error characteristics of fiber optic transmission systems*, Delayed contribution 21, CCITT Study Group XVIII, San Diego, 23 jan 1989.

[10] *DS-1 Facility performance: optical fiber as a long haul technology*, Proceedings of ICC-877, pp. 646-649.

[11] *Impact of transmission error on ATM network performance*, K.A. Turban and A. Engel.

[12] *Communication Quality analysis for ATM Networks*, K. Takahashi et al., IEEE ICC 1989, Vol. 1, pp. 423-427.

[13] *Quality of service requirements in ATM-based B-ISDNs*, M.E. Anagnostou et al., Computer communications, Vol. 14, no.4, may 1991, pp. 197-204.

[14] CCITT Recommendation X.212.

[15] *Requirements for signal transfer protocol over ATM network*, Delayed contribution D.1410, CCITT SG XVIII/8, Geneva, June 1991.

[16] *VP switch in an ATM network*, Delayed contribution D.1481, CCITT SG XVIII/8, Geneva, June 1991.

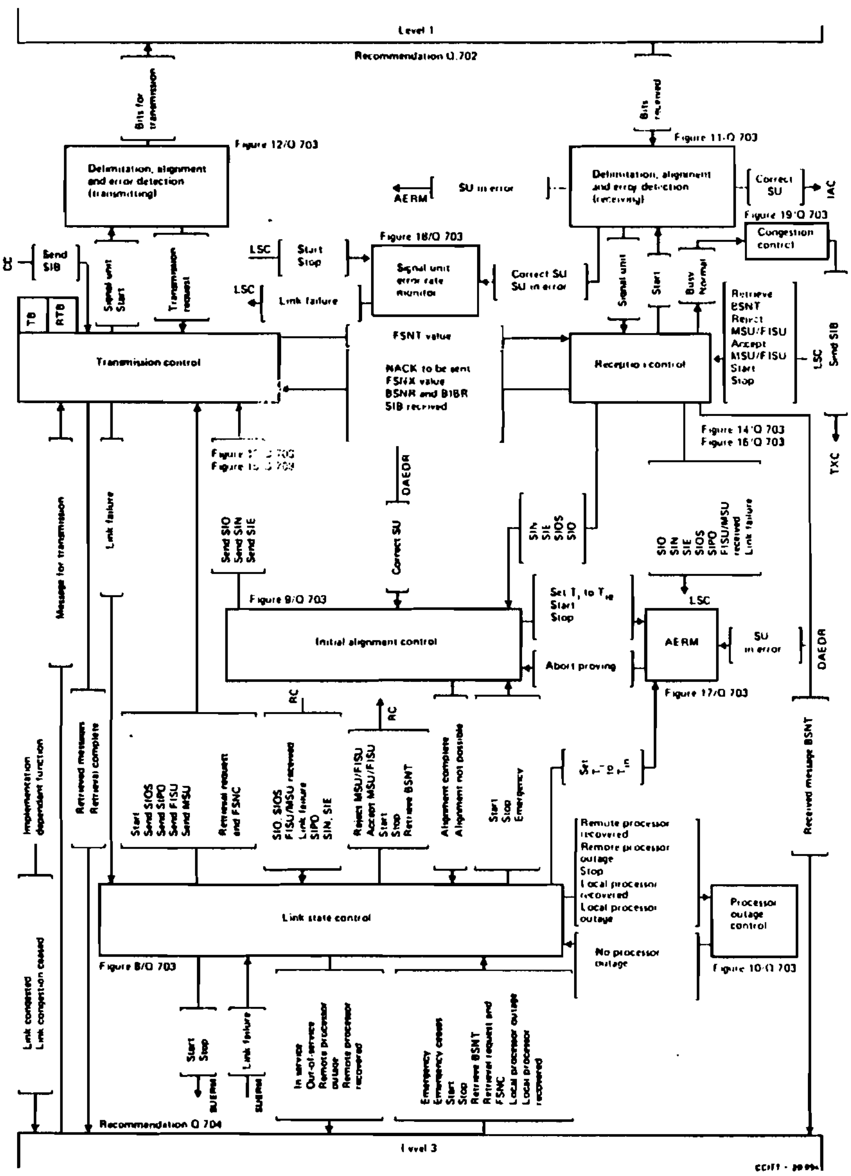
[17] *B-ISDN metasignalling Baseline Document*, CCITT SG XI/4, Geneva, April 1991.

[18] *Telecommunications, Protocols and Design*, J.D. Spragins, Addison-Wesley Publishing Company, Inc., 1991

[19] *OSI-protocollen, lagen 1 t/m 4*, O.B.P. Rikkert de Koe, Kluwer technische boeken b.v., 1990

[20] *Liaison Statement to WP XVIII/4 and XVIII/8, Intergration of Satellites in the B-ISDN*, Experts Group Meeting, 30 April - 3 May 1991, Baltimore, Maryland.

Appendix A: MTP messages



Note 1 – Abbreviated message names have been used in this diagram (i.e. origin – destination codes are omitted)

Note 2 – See the abbreviations and timers used in this figure in § 12.

FIGURE 7/Q.703

Level 2 – Functional block diagram

Figure A.1: Interaction between MTP level 2 and 3

Appendix B: HDLC

This Appendix describes a summary of the HDLC protocol [18].

The most widely publicized and implemented current generation DLC protocols are bit-oriented protocols. HDLC falls into this class of protocols.

In HDLC the frames are exchanged between a primary and a secondary station. Frames from primary to secondary are called commands and frames from secondary to primary are called responses. In the original protocol the only allowable configuration were the unbalanced point-to-point configuration and the unbalanced multipoint configuration. Political pressures, however, led to allowing a configuration with two combined stations (stations containing both a primary and a secondary) on a point-to-point link.

The normal format of HDLC frames is given in the figure below:

Field	Flag	Address	Control	Information	FCS	Flag
Length (octets)	1	1	1		2	1

The Flag field contains an eight-bit character, 01111110. This bit pattern is not allowed to occur in a transmitted frame between flags (bit stuffing ensures this). The Address field always contains the address of the secondary station. The Control field identifies the type of frame and contains frame sequence counts. The Information field contains user data. The FCS field contains a frame check sequence which checks the contents of the address, control and information field.

HDLC has three different formats: Information, Supervisory and Unnumbered.

Information frames are used for normal data transfer. Piggy-backing of acknowledgements on information frames and acknowledgement of several frames at once are allowed.

Supervisory frames are used for positive and negative acknowledgements, if no user data are available for return frames, as well as for other purposes. There are four types of supervisory frames, these are: Receive Ready (RR), Receive Not Ready (RNR), Reject (REJ) and Selective Reject (SREJ). An RR frame is a general purpose frame most often used for acknowledgement. RNR is used for indicating the station is temporarily not ready to receive information frames. REJ is a negative acknowledgement and requests recovery by go-back-N retransmission. SREJ is also a negative acknowledgement but only requests retransmission of the negatively acknowledged frame.

Unnumbered frames are used for housekeeping purposes, including link startup and shutdown, specifying modes of operation.