

Services and Protocols for Advanced Networks (SPAN); Relationship between IP and telecommunication networks



Reference

DEG/SPAN-140103

Keywords

IP, network

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

Individual copies of the present document can be downloaded from:

<http://www.etsi.org>

The present document may be made available in more than one electronic version or in print. In any case of existing or perceived difference in contents between such versions, the reference version is the Portable Document Format (PDF). In case of dispute, the reference shall be the printing on ETSI printers of the PDF version kept on a specific network drive within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at <http://www.etsi.org/tb/status/>

If you find errors in the present document, send your comment to:
editor@etsi.fr

Copyright Notification

No part may be reproduced except as authorized by written permission.
The copyright and the foregoing restriction extend to reproduction in all media.

© European Telecommunications Standards Institute 2001.
All rights reserved.

Contents

Intellectual Property Rights	5
Foreword.....	5
Introduction.....	5
1 Scope.....	7
2 References.....	7
3 Definitions and abbreviations.....	9
3.1 Definitions	9
3.2 Abbreviations.....	13
4 Conceptual framework.....	14
4.1 Traffic types.....	14
4.2 Protocol layers vs. fundamental protocol functions.....	14
4.3 Fundamental network functions.....	14
4.3.1 Access	15
4.3.2 Traffic management.....	15
4.3.3 Naming, addressing and routing	15
4.3.4 Stream transport.....	15
4.4 Generic network model.....	16
4.5 Different networking paradigms	17
4.6 The problems	18
5 Access functions.....	19
5.1 Key issues.....	19
5.2 The use of digital subscriber lines and ISDN for access.....	19
5.3 The use of Ethernet (IEEE 802.3 CDMA/CD) as an access protocol.....	19
5.4 The use of ATM as an access protocol	19
5.5 The use of DTM as an access protocol	20
6 An analysis of IP interworking with other protocols.....	20
6.1 Basic assumptions	20
6.2 Problems with traffic management under packet switching.....	20
6.3 Interworking cases; IP to fibre	21
6.4 IP over ATM.....	23
6.4.1 IP/AAL5/ ATM (Cases 2A, 2B, 2C, 2D, 2E)	23
6.4.2 IP/LLC/SNAP/AAL5/ATM over SDH (Cases 2B + 31).....	24
6.4.3 IP/LLC/SNAP/AAL5/ATM over fibre (Cases 2B + 33).....	25
6.4.4 IP/LANE/AAL5/ATM (Case 2D).....	25
6.4.5 IP/PPP/AAL5/ATM (Case 2E).....	26
6.5 IP over PDH (Case 1).....	26
6.6 IP/PPP over Frame Relay (Case 3).....	26
6.7 IP over DTM (Case 8).....	27
6.8 Frame Relay over ATM (Case 22).....	27
6.9 Ethernet over DTM (Case 11).....	28
7 Stream transport in an integrated transport network.....	29
7.1 Basic assumptions	29
7.2 IP over SDH.....	30
7.2.1 IP/PPP/HDLC over SDH or SONET (Case 4).....	30
7.3 IP over SRP over SDH (Case 5).....	31
7.4 IP over Gigabit Ethernet over WDM transponder/fibre (Cases 7 + 12).....	32
7.5 IP over PPP over SDL over fibre (Case 9).....	33

8	Comparison of available bandwidth	34
Annex A (informative):	Bibliography.....	35
History		36

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (<http://www.etsi.org/ipr>).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This ETSI Guide (EG) has been produced by ETSI Technical Committee Services and Protocols for Advanced Networks (SPAN).

Introduction

ITU-T SG 13 has defined a new project, IP and Telecommunications Networks Inter-relationships. This project is managed within WP 1/13. It has been decided within ETSI WG SPAN 8 to initiate a supporting work item with the working title as above.

The problem the present document is addressing is how to achieve the most efficient integration of traffic types and existing protocols which have been designed and used in different technical cultures.

IP and related protocols designed by IETF have their roots in the need to transport files without guarantees of timely packet delivery which might result in packet loss. ATM and SDH or SONET have been designed by ITU-T and ATM Forum have their roots in the need to transport voice and telephony based services (fax) through circuit-switching based networks under very controlled circumstances. DTM is a new protocol system aiming at QoS management of real-time flows. In a situation where IP and its data format has been a common denominator for data communication access, and the volume of data traffic is overtaking voice communication, the interworking between circuit-switched and packet-switched networks has come under heavy debate.

The approach of the present document is to return to fundamental problems and principles which may serve as a basis for a non-biased discussion on the most effective protocol system that can integrate traffic types of completely different nature.

As can be inferred from the definitions of "IP network" as well as "telecommunication network" this classification is no longer very much technically relevant. There is no clear distinction any more. The only possible technical ground for an analysis of the relationships between networks of different historical origin, is to relate the specifications of those protocols to a canonical model of networking. The major source of this canonical model is ITU-T Recommendation G.803 [1] and ETS 300 299 [2].

The technically interesting borderline goes between real-time CBR bit-loss tolerant services and non-real time VBR non bit-loss tolerant services. These kinds of services are so different with respect to their requirements on the underlying network for routing/switching and transmission, that the commercial necessity to build and run one integrated services network becomes a very demanding technical challenge. This is what the network design problem is all about. Various techniques for traffic analysis, service classification, traffic policing, flows control, network management etc all aim at handling each kind of service in the most economical way in one integrated services network.

The historical background has led us to the current situation where many independent technical bodies are engaged in the production of specifications and standards for systems that have to interwork very tightly in the current commercial networking business. None of these bodies can now produce final technical specifications without checking the interrelationships with specifications from other bodies. The networking community is therefore now occupied with a web of many formal as well as informal technical liaisons.

It is in this context we shall see the present document. For ITU-T as well as ETSI, it has been found necessary and desirable to fully recognize the existence and justification of informal and voluntary technical standardization bodies such as IETF, IEEE, ATM Forum and Multiservice Switching Forum. The important aspect for the community of users of standards, manufacturers as well as operators, is to avoid a situation where political prestige or particular commercial interests are delaying or misguiding the standardization process.

The focus of the present document is therefore to identify those areas where the technical interfaces between standards and specifications from different bodies are causing problems for the community of standards users. These results can then be used in open and fair discussions between independent bodies to find which body is the best forum for solving these problems.

1 Scope

The following aspects of the relationship between IP networks and telecommunication networks are included:

- 1) access to IP networks using telecommunication facilities;
- 2) interworking between IP networks and telecommunication networks (e.g. gateway functions between IP and other protocols on layer 3);
- 3) transport mechanisms for IP networks (e.g. IP over other protocols).

The following aspects, which are within the scope of the ITU-T work, are not included:

- 1) application interworking;
- 2) the integrated use of signalling in IP as well as traditional networks;
- 3) analysis of future IP protocols;
- 4) access techniques for telecommunication networks.

NOTE: The ETSI project TIPHON is the lead body for Multimedia over IP.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication and/or edition number or version number) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies.

- [1] ITU-T Recommendation G.803 (2000): "Architecture of transport networks based on the synchronous digital hierarchy (SDH)".
- [2] ETSI ETS 300 299 (1997): "Broadband Integrated Services Digital Network (B-ISDN); Cell based user network access for 155 520 kbit/s and 622 080 kbit/s; Physical layer interfaces for B-ISDN applications".
- [3] ITU-T Recommendation I.363 (1993): "B-ISDN ATM adaptation layer (AAL) specification".
- [4] ITU-T Recommendation I.363.5 (1996): "B-ISDN ATM Adaptation Layer specification: Type 5 AAL".
- [5] IETF RFC 1483 (1993): "Multiprotocol Encapsulation over ATM Adaptation Layer 5".
- [6] IETF RFC 2225(1998): "Classical IP and ARP over ATM".
- [7] IETF RFC 1755 (1995): "ATM Signalling Support for IP over ATM".
- [8] IETF RFC 2364 (1998): "PPP Over AAL5".
- [9] IETF RFC 1973 (1996): "PPP in Frame Relay".
- [10] IETF RFC 2615 (1999): "PPP over SONET/SDH".
- [11] ETSI ETS 300 300 (1997): "Broadband Integrated Services Digital Network (B-ISDN); Synchronous Digital Hierarchy (SDH) based user network access; Physical layer User Network Interfaces (UNI) for 155 520 kbit/s and 622 080 kbit/s Asynchronous Transfer Mode (ATM) B-ISDN applications".

- [12] ITU-T Recommendation I.432.1 (1999): "B-ISDN user-network interface - Physical layer specification: General characteristics".
- [13] ITU-T Recommendation I.432.2 (1999): "B-ISDN user-network interface - Physical layer specification: 155 520 kbit/s and 622 080 kbit/s operation".
- [14] ATM Forum af-lane-0084.000 (1997): "LANE v2.0 LUNI Interface".
- [15] ATM Forum af-lane-0112.000 (1999): "LAN Emulation over ATM Version 2 - LNNI Specification".
- [16] IETF RFC 1661 (1994): "The Point-to-Point Protocol (PPP)".
- [17] IEEE 802.3z: "Media Access Control (MAC) Parameters, Physical Layer, Repeater and Management Parameters for 1 000Mb/s Operation".
- [18] IETF RFC 2892 (2000): "The Cisco SRP MAC Layer Protocol".
- [19] ITU-T Recommendation G.707/Y.1322 (2000): "Network node interface for the synchronous digital hierarchy (SDH)".
- [20] ITU-T Recommendation G.983.1 (1998): "Broadband optical access systems based on Passive Optical Networks (PON)".
- [21] ITU-T Recommendation I.113 (1997): "Vocabulary of terms for broadband aspects of ISDN".
- [22] ITU-T Recommendation E.164 (1997): "The international public telecommunication numbering plan".
- [23] IETF RFC 2105 (1997): "Cisco Systems' Tag Switching Architecture Overview".
- [24] ATM Forum af-uni-0010.002 (1994): "ATM User-Network Interface Specification V3.1".
- [25] ATM Forum af-phy-0128.000 (1999): "622 and 2 488 Mbit/s Cell-Based Physical Layer".
- [26] ATM Forum af-phy-0133.000 (1999): "2,4 Gbps Physical Layer Specification".
- [27] ATM Forum af-phy-0046.000 (1996): "622,08 Mbps Physical Layer".
- [28] ITU-T Recommendation I.363.2 (2000): "B-ISDN ATM Adaptation Layer (AAL) type 2 specification".
- [29] IEEE 802.1Q (1998): "IEEE Standard for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks".
- [30] ISO/IEC 15802-3 (1998): "Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Common specifications - Part 3: Media Access Control (MAC) Bridges".
- [31] IETF RFC 1662 (1994): "PPP in HDLC-like Framing".

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the following terms and definitions apply:

access: function that enables connections from an end user equipment upstream as well as downstream

NOTE 1: The two directions of the communication can take different routes.

aggregate stream: aggregation of many individual streams

NOTE 2: Depending on the type of the component streams, we have real-time (aggregate) streams or non real-time (aggregate) streams.

best-effort relationship: particular kind of connection (relationship) between two nodes A and B for which no commitment exists, but where it is possible that a datagram accepted at node A will arrive at node B

NOTE 3: However, there is no guarantee that the datagram will arrive at node B.

NOTE 4: Connection (including circuit and best-effort relationship) is the only possible relationships for nodes that can be related.

cell: packet of fixed length (see ITU-T Recommendation I.113 [21])

characteristic information: those parts of a format definition of the basic traffic entity of a layer network which is transported unchanged across a connection or circuit

NOTE 5: Characteristic information is always defined in relation to a particular layer network.

For example, characteristic information on layer 2 may not be characteristic information on layer 3, since it can be changed when a traffic entity instance is moving across a network node.

For example, in IP a characteristic information is the information field of the IP packet but not its header which can be changed as the packet crosses the network.

circuit: relationship between two not necessarily adjacent nodes A and B such that there exists a connection from A to B with allocated bandwidth

NOTE 6: Thus a circuit is always a connection, by its definition at network layer.

EXAMPLE 1: Given connections A-B and B-C, there exists a circuit A-C that fulfils the minimal service requirements fulfilled by A-B and B-C.

connection: relationship between two endpoints A and B such that a flow can be transported from A to B under fulfilment of given service requirements. Only the endpoints are required to know of the connection while it does not exclude that the network or part of the network knows of said connection

NOTE 7: There is a very large number of possible distinct connections, based on the powerset of possible connection parameter value sets. (I.e. all parameter value combinations).

NOTE 8: A flow may be transported within a connection.

connection admission control: function that assesses whether there is sufficient resource to admit a connection across a subnetwork

NOTE 9: A connectionless network does not have a connection admission control function.

connection control: function that changes parameter values for a connection

NOTE 10: The scope a connection control function can be a single connection (link layer) or a single circuit (network layer). This function includes setting the rules for congestion handling.

connectionless network: network layer network with respect to nodes A and B such that when sending data from A, the data is not explicitly routed to B

NOTE 11: In order for a network data unit (e.g. IP packet) to find its destination B, a subnet has to set up ad-hoc link layer connections step-by-step and use them on a best-effort basis until the data unit has reached B. The network is in many respects non-deterministic:

- it is not guaranteed that B exists at run time;
- it is not guaranteed that there will exist a path from A to B;
- it is not guaranteed that intermediate nodes will take care of the data unit when it arrives;
- thus there is no guarantee at all the data unit will arrive to B.

connection oriented network: network layer network with respect to nodes A and B such that when sending data from A, some relationship exists and is invariant during transmission

NOTE 12: Such invariant may be very strong or very weak.

The strongest invariant is a reserved resource from A to B.

A weak invariant may be a commitment (such as in TCP) that only ensures that data eventually will reach B when transmitted from A.

Thus, a network with no invariants at all is a connectionless network.

content integrity: relationship for a connection A-B such that bits sent from node A are received unchanged at node B

NOTE 13: This relationship need only be maintained when data is in transit from A to B.

datagram: datagram is a packet with full address information enabling it to be routed to the endpoint without further information

Datagram control: functions that control the integrity of datagrams

NOTE 14: Checksums may be used to control content integrity. Timestamps may control timing integrity.

file: in the context of the present document a traffic type denoting finite flows for which content integrity is of importance

NOTE 15: File size is normally known before transmission starts. The knowledge of size may have implications for the connection.

file transport connection: connection that in some way is capable of transporting traffic type File

NOTE 16: The major requirement for this kind of connection is to support content integrity.

flow: unidirectional stream of packets that are sent from a particular source to a particular destination (unicast or multicast) address and any logical handling policy they may require

NOTE 17: A flow is a service instance, that is managed as a single entity.

flow control: capabilities for control of a flow

NOTE 18: Not to be mixed up with connection control, which is controlling the resource used by a flow.

flow transport: transport of a flow through a connection

NOTE 19: In this case, it is assumed that the flow is controlled individually. It may have a dedicated connection, but this is not necessary; it can also be using a shared resource.

frame: sequence of bits forming a delimitation of contained data

NOTE 20: In RFC 1661 [16] defined as: The unit of transmission at the data link layer. A frame may include a header and/or a trailer, along with some number of units of data.

integrated connection: connection that supports at least two traffic types

NOTE 21: A connection that supports a traffic type with certain requirements i.e. has invariants of a certain strength, can also support all traffic types which logically demand invariants that are weaker.

interactive real-time stream: real-time stream related to an interactive application

NOTE 22: An interactive real-time stream must use a connection that fulfils the constraints for non-interactive streams plus minimum round-trip delay requirements.

interactive real-time transport connection: connection that in some way is capable of transporting traffic type interactive real-time flow

NOTE 23: From a timing integrity (delay) point of view, this is the most qualified connection.

IP network: network that is accommodating IP and protocols related to IP

NOTE 24: Pragmatically networks defined by IETF.

layer network: connection control and flow control system capable of handling traffic instances which have the same characteristic information

NOTE 25: Usually a layer network has a related addressing scheme.

NOTE 26: IP and ATM are examples of layer networks, capable of handling IP and ATM flows respectively.

NOTE 27: Each layer network can handle at least one traffic type and all of its instances. For example, a correct IP layer network shall be able to handle all possible IP flows.

message: traffic type where the instances are datagrams related to events in a controlled system

NOTE 28: Messages are usually small, less than 1 kbit/s.

NOTE 29: Signals in the control plane are typical messages.

message transport network: network that in some way is capable of transporting traffic type Message

NOTE 30: A signalling network is a typical message transport network,

multi-cast connection: set of connections such that each connection has invariants that fulfil requirements between one sender node and a set of receiver nodes

EXAMPLE 2: The set of connections {A-B, A-C, ..., A-N} is a multi-cast connection.

NOTE 31: Some connections in the set may be stronger than others; there is no need for them to be equally strong.

network: set of connectable nodes accommodating functions of at least one protocol layer

NOTE 32: The set of connectable nodes accommodating protocol layer X is a X-layer network.

For example, a set of connectable nodes accommodating IP functions is an IP-layer network.

NOTE 33: A multi-layer network is a set of overlaid layer networks, e.g. an IP layer network over an ATM layer network.

non-interactive real-time flow: real-time flow which is serving a non-interactive application

NOTE 34: E.g. a video-on-demand transmission from repository to video player for direct view.

non-interactive real-time stream: collection of non-interactive real-time flows

NOTE 35: E.g. several TV channels transported over one physical medium in an interleaved mode.

non-interactive real-time transport connection: connection that in some way is capable of transporting traffic type non-interactive real-time flow

non-real-time flow: flow which is serving a non real-time application

NOTE 36: E.g. a file transfer instance.

non real-time stream: collection of non-real-time flows
e.g. a set of interleaved ATM cells transporting several files in parallel.

packet: logical grouping of bits having variable length with control information

NOTE 37: In RFC 1661 [16] defined as: The basic unit of encapsulation, which is passed across the interface between the network layer and the data link layer. A packet is usually mapped to a frame; the exceptions are when data link layer fragmentation is being performed, or when multiple packets are incorporated into a single frame.

real-time flow: flow that is serving a real-time demanding application, where the time position of each piece of information in the flow is significant

real-time stream: collection of real-time flows e.g. a set of interleaved telephone calls transported in simplex mode.

NOTE 38: Real-time streams may recursively be multiplexed into higher-order real-time streams. E.g. SDH/SONET. It must be supported by a protocol that is synchronous or plesiochronous.

There are two kinds of real-time streams: interactive real-time streams and non-interactive real-time streams.

real-time transport connection: connection that in some way is capable of transporting traffic type Real-time flow

NOTE 39: The major requirement for this connection is to support timing integrity.

NOTE 40: There are two kinds of requirements: Non-interactive real-time transport connection and interactive real-time transport connection.

stream: reoccurring transmission of data bound by some timing requirements

NOTE 41: Streams may be real-time or non-real-time. Streams are either sequential or parallel over some particular resource.

stream transport: function that enables transmission of streams through a connection

NOTE 42: Stream creation includes segmentation, adding of control data and other adaptations to the physical layer protocol.

telecommunications network: network that is mainly accommodating connection-oriented protocols for telecommunication services

NOTE 43: Examples: PSTN, ISDN, ATM, PDH, SDH or SONET.

NOTE 44: Pragmatically, networks defined by ITU-T or telecom related regional bodies such as ETSI.

timing integrity: relationship for a connection A-B such a bit received at node A was sent by node A not earlier than a defined time unit

NOTE 45: This relationship need only be maintained when data is in transit from A to B.

NOTE 46: This relationship, as it is defined above, does not preclude loss of bits in transit. It says that if a bit arrives at B, it is delayed within a certain limit. Thus if all bits shall arrive with a bounded delay, both content and timing integrity must be demanded.

traffic type: traffic for which the same requirements, and thus the same rules and functions apply for all flows of the type

NOTE 47: There are four traffic types: message, file, real-time non-interactive and real-time interactive.

traffic type instance: flow of a certain type

NOTE 48: IP packet flows and ATM cell flows are examples of traffic type instances that can be of type message, file, real-time non-interactive or real-time interactive depending on the served application.

virtual circuit: a logically emulated circuit

NOTE 49: The term "virtual" indicates that the connection is of an abstract nature.

3.2 Abbreviations

For the purposes of the present document, the following abbreviations apply:

AAL	ATM Adaptation layer	ATM Forum
ABR	Available bit rate	ATM Forum
AP	Access point (to the access network)	
APS	Automatic protection switching	
ARP	Address resolution protocol	IETF
ATM	Asynchronous transfer mode	ATM Forum
CBR	Constant bit rate	ATM Forum
CSMA/CD	Carrier sense multiple access with collision detection	IEEE 802
DNS	Domain name server	IETF
DSS	Distributed Sample Scramble	ATM Forum
DTM	Dynamic synchronous transfer mode	
DWDM	Dense WDM	
GBE	Gigabit Ethernet	
HDLC	High-level data link control	IEEE
IEEE	Institute of Electrical and Electronics Engineers	
IETF	Internet Engineering Task Force	
IP	Internet protocol	IETF
ITU-T	International Telecommunication Union, Telecommunication Standardization Sector	
LAN	Local area network	IEEE
LAPD	Link access protocol, D-channel	IEEE
MAC	Medium access control	IEEE
OChP	To be defined (by TeleDanmark)	
OMSP	To be defined (by TeleDanmark)	
OSI	Open systems interconnection	
PDH	Plesiochronous digital hierarchy	
PPP	Point-to-point protocol	IETF
PVC	Permanent virtual channel	ATM Forum
QoS	Quality of service	
SDH	Synchronous digital hierarchy	
SDL	Simple Data Link	
SONET	Synchronous optical network	
SRP	Spatial reuse protocol	
TCP	Transmission control protocol	IETF
UBR	Unspecified bit rate	ATM Forum
UDP	User Datagram Protocol	IETF
UNI	User-Network interface	ATM Forum
VBR	Variable bit rate	ATM Forum
VBR-rt	Real-time VBR	ATM Forum
VPC	Virtual path connection	ATM Forum
WDM	Wavelength division multiplexing	

4 Conceptual framework

4.1 Traffic types

We start with the assumption that the traffic to be handled by a network is constituted by the following four main categories, here called *traffic types* (see note):

- 1) messages;
- 2) files;
- 3) non-interactive real-time stream;
- 4) interactive real-time stream.

NOTE: Some sources call them "traffic entities". However, it should be better to draw an analogy to the theory of data types. A "traffic type" is thus a template for traffic of a certain characteristic for which the same requirements apply to all instances of the type. Thus infinite subtyping is theoretically possible.

The characteristic demands for Message are that:

- messages are transported error free;
- messages always arrive;
- messages arrive in defined order;
- the latency is guaranteed to be within certain protocol defined limits, i.e. before timeout.

Message type demands are undefined with respect to bit arrival rate. However, due to the small size, this is not significant from a traffic management point of view, as a message normally is contained in one single packet or frame.

The demand for File is normally relaxed with respect to latency and bit rate. The reason is that a file is transported from a buffer in one place to a buffer in another place. It is thus neither produced integrated with transmission, nor consumed integrated with transmission.

The demand for Real-time-flow can be relaxed with respect to content integrity. This is particularly the case when the flow is consumed by a human being at arrival time. (The human being cannot notice minor bit losses anyway). Bit loss relaxation is part of the service specification, and should be carefully considered, since the cost implications of too strict bit loss requirements are significant.

4.2 Protocol layers vs. fundamental protocol functions

The problem for the telecom operator business today is that several protocols are clashing, and that there are too many ideas how these clashes can be resolved. The proper approach to solve that problem is to revert to basic architectural principles. Whatever the names of protocols, and borderlines between layers, there are fundamental functions that have to be implemented in protocols depending on service requirements.

4.3 Fundamental network functions

There are essentially only four fundamental functional groups of a network, regardless it is called a telecom network or a datacom network, as long as the traffic types mentioned above are involved:

- 1) access;
- 2) traffic management;
- 3) naming, addressing and routing;
- 4) transport.

4.3.1 Access

This functional group is related to the interworking between an end point device (producer or data repository) at the access point (AP) and the first network element performing traffic management (edge device). Between AP and edge device there is no traffic management and simplest possible individual transmission. The traffic is adapted to the common network service in the SCP. The communication link is direct and always open from AP to edge device. The function is just to bridge the distance from point to point.

4.3.2 Traffic management

This functional group is necessary due to the fact that a shared resource such as a public network has to serve several users with traffic requirements that are varying over space and time (i.e. if there were a permanent connection between any sender and receiver then there would be no need for traffic management or data transformation).

Among traffic management functions we find (reference: GIGA 98):

- traffic contracting (traffic description, QoS objectives, conformance definition);
- connection admission control;
- conformance monitoring;
- queuing and scheduling;
- flow control;
- congestion control.

Traffic management functions are related to certain protocol layers depending of the scope of the management. In general, if the management scope includes all network nodes from edge device to edge device (i.e. a shared resource network), the functions belong to the network layer (OSI layer 3). If the scope includes two nodes directly connected via a physical connection then the functions belong to the data link layer (OSI layer 2).

4.3.3 Naming, addressing and routing

The naming, addressing and routing are important aspects of large scale networks and their engineering. The design and use of the naming and addressing system have a big influence on network behaviour and the management of the network.

Currently two major systems are in use, these are ITU-T Recommendation E.164 [22] (with enhancements) and the Internet IP and DNS systems.

Both systems distinguish between a user identifier and a location identifier. The location identifier is usable for routing throughout the network. The user identifier is mapped into a location identifier by the network prior to further distribution through the network.

4.3.4 Stream transport

This functional group is needed as a generic resource that very cost-effectively can transport bits representing data between edge devices.

Note that stream transport of bits as a problem area can involve very complicated functions (e.g. fault management) which are irrelevant from a traffic management point of view, and thus excluded from the present document.

In general, stream transport functions are related to the physical layer (OSI layer 1).

Note that from one layer network's point of view there is a physical layer network, although this layer network may be abstract and the service is performed by a more or less complete stack of protocol layer networks (e.g. when an ATM network is utilizing an SDH network as physical layer).

4.4 Generic network model

This figure is intended to cover the major functional divisions of a public network:

- 1) access;
- 2) aggregation and edge functions;
- 3) transport.

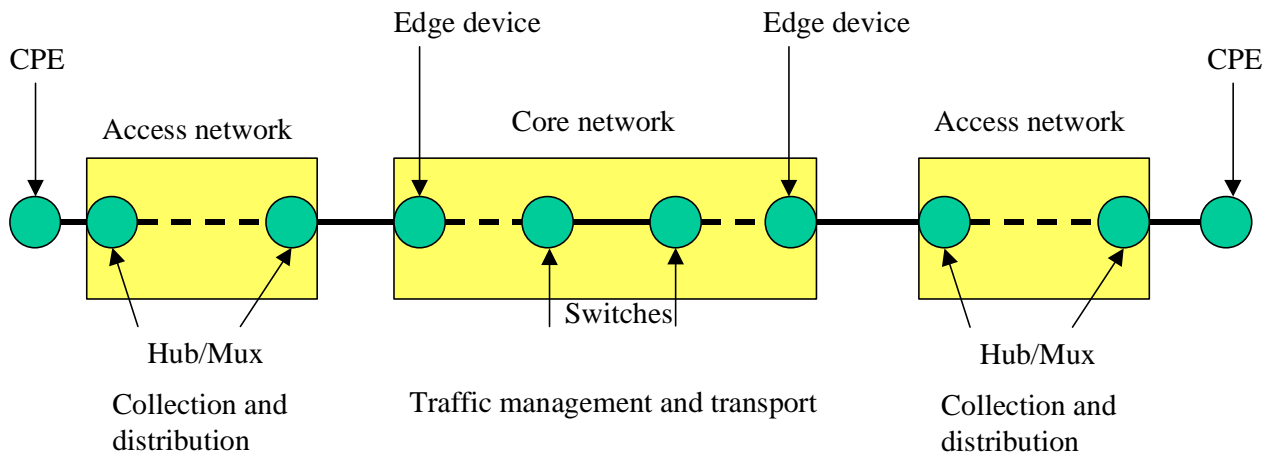


Figure 1: Generic network model

Seen from left to right as a simplex traffic example, the traffic type instance enters an AP. Between the AP and the edge device in the public network there is a direct line. This direct line can pass a system of multiplexers in the access network. The edge device in the core network can be a router/switch, or there is an intermediate network that is performing the role of aggregating traffic up to edge routers/switches. Finally, there is a core network that is accepting routed traffic and transmitting it through the core network, via connections of various strength (in reality, for performance reasons, no network is using pure hop-by-hop routing any more). When the traffic leaves the core network, the process is reversed.

During its way through the network, the traffic type instance is subject to the various network functions mentioned above. The general problem for the operator is to manage the functions such that service requirements for many different traffic types are fulfilled at the lowest possible cost.

From a protocol and layered network point of view, we can define a generic layered network model as follows:

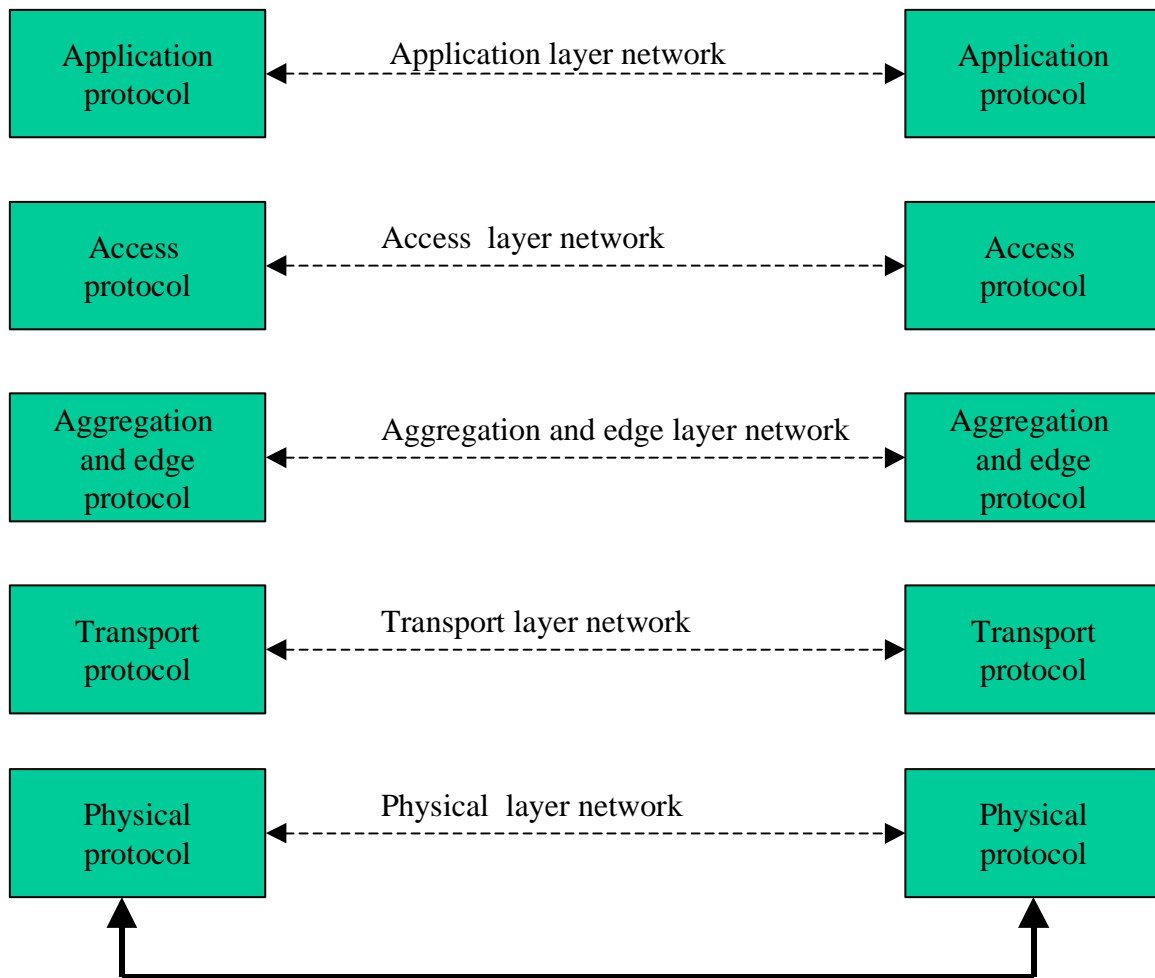


Figure 2: Generic layered network model

4.5 Different networking paradigms

The title of the present document indicates that there are two kinds of networks: "IP networks" and "Telecommunication networks". This distinction is an economic reality, although from a technical point of view, there is no justification whatsoever. The distinction is gradually becoming blurred, on the technical as well as the commercial level.

In order to start the analysis, it is nevertheless necessary to recall the ad-hoc definitions for the present document:

Datacommunications network	A network that is accommodating datacommunication protocols (like IP).
Telecommunications network	A network that is mainly accommodating connection-oriented protocols for telecommunication services.

Two cases of interworking must be distinguished, service interworking and network interworking.

The taxonomy that should be more useful for analysis should be based on traffic types and protocols. We can either talk about networks optimized for – or at least capable of - handling certain traffic types, such as Message transport network, File transport network, Real-time flow transport network and Integrated transport network. (See definitions). Or we can talk about particular layer networks, such as IP layer network, ATM layer network etc.

In general, concerning layer networks, we can adopt the notation that a protocol stack from top to bottom defines a particular multi-layer multiprotocol network that has certain capability characteristics. For example the common IP/MPLS/AAL5/ATM/SDH multi-layer network.

In the older days, when a data network was dedicated to files and a telecommunication network was dedicated to voice, then the two paradigms of connection-less and connection-oriented were clearly separated. ISDN was already, on the other hand, the first approach to integrate network functions to serve different traffic types. Switched Ethernet was another example of merged approaches. ATM is clearly, under the name of B-ISDN, an effort to integrate network functions to serve different traffic types. In IETF, the name Diff-Serv also point at various efforts to introduce new connection-oriented functions into an essentially connection-less networking paradigm.

4.6 The problems

One problem that now is facing the networking community, based on the terminology introduced in the definitions clause, is the following:

Problem 1 (Integrated network approach)

Given a requirements profile defined as a mixture of traffic types to be served, is there a multiprotocol network available that will fulfil the requirements at acceptable cost?

Or, if there is no single multiprotocol network available that will fulfil all requirements:

Problem 2 (Several optimized networks approach)

Given a requirements profile defined as a mixture of traffic types to be served, is there a set of multiprotocol networks available that will fulfil the requirements at acceptable cost?

In the latter case, the problem for the network operator is that several distinct networks must be acquired, managed and used in parallel. This may lead to increased complexity as concerns network management. But this is not generally so. Some would argue that two or even several simple and optimized networks would lead to easier operation, better QoS for all traffic types and higher net profit.

For the present document the problem is not only to analyse existing protocols, but also try to find a new solution by integrating the two historically different paradigms:

Problem 3 (New integrated network approach)

Given a requirements profile defined as a mixture of traffic types to be served, is it possible to define a new multiprotocol network that will fulfil the requirements at acceptable cost?

5 Access functions

5.1 Key issues

The major requirement for an access related protocol is that it is ubiquitous. The major reason why IP has become so widely used is that the web-browsing service is easy to use and thus have reached a wide deployment.

Since then, IP (and in particular IPv4) has become the common denominator to most underlying network services, even though it may not be perfect for this purpose. ATM as access protocol is in doubt, not primarily because of its technical characteristics, but because IP has gained such popularity. Ethernet (CSMA/CD) has always been a popular access protocol for LANs.

The key issues with respect to access for all traffic types are:

- 1) the use of PPP for access (over 90 % of all real life cases);
- 2) the use of Ethernet as an access protocol (major LAN protocol);
- 3) the use of token ring as an access protocol (less than 5 %);
- 4) the use of ATM as an access protocol (used in conjunction with ADSL);
- 5) the use of DTM as an access protocol (possible solution).

These issues will be commented one by one below.

5.2 The use of digital subscriber lines and ISDN for access

Due to the fact that there are enormous investments made into copper lines from distribution nodes in the telecom network, there is a desire from incumbent operators to utilize these investments also for higher-bandwidth demanding services. With the use of new protocols and access devices, it is possible to offer asymmetric access up to bandwidth in the order of 2 - 5 Mbps.

There are technical problems related to the fact that twisted pair cables may not be sufficiently shielded to avoid mutual interference in high frequencies.

5.3 The use of Ethernet (IEEE 802.3 CDMA/CD) as an access protocol

Ethernet is essentially a local access and transmission system. It is not very well suited for real time applications, since the access to the shared transmission medium is non-deterministic, and a user will not be allowed to monopolize the medium too long. However, if a switch instead of a shared access medium is used, the characteristic feature of collision detect is lost, and the protocol will be similar to any other using a multi-channel switch. Switched Ethernet can also serve for aggregation.

Gigabit Ethernet will most likely be used for stream transport point-to-point between LANs over private networks using dark fibre where QoS and resource sharing is not an issue.

5.4 The use of ATM as an access protocol

ATM is used as an access protocol in ADSL systems and also in LMDS wireless systems (see [14]).

5.5 The use of DTM as an access protocol

The default DTM topology is a dual ring. In those cases where nodes on the ring are stretched into customer premises, the DTM ring can replace an Ethernet LAN. It is however, only in very special circumstances cost/benefit justified to use a DTM LAN instead of an Ethernet LAN. In all other cases, the LAN in the customer premises will be Ethernet and the Ethernet LAN will interface directly to an access node for DTM in the curb or the basement. The conclusion is that DTM will not find in the market a position as a major access protocol. On the other hand, DTM due to its design based on fibre rings will be a good choice for aggregation networks.

6 An analysis of IP interworking with other protocols

6.1 Basic assumptions

Concerning the dominant network structure, we can make the following assumptions, which seems now to be consensus within the telecom/datacom business:

- 1) for economical reasons, there should be one and only one integrated transport network for all traffic types. However, attached to such a network, many options for access protocols will be available. In principle, each traffic type and each bit rate category will be offered a particular access protocol. Access options will proliferate and be very commercially driven, while the core network will be very standardized;
- 2) at the edge of core networks there will be large service control points designed as hybrid routers/switches that will analyse requests and decide the best path and general treatment of a service instance. These edge devices will have access to address tables as well as relevant requirements information. The problem for edge devices is to decide what should be the strength of the resource reservation of the connection given that the service requirements are known. Note that IETF DiffServ and Class of Service as well as ATM Forum service categories are ways of informing the edge devices how to manage traffic through the core network;
- 3) core networks will have a tendency to be designed towards connection-oriented (switched) to serve large files and all real-time flows. This design principle will enable reliable transfer of large files as well as QoS management for real-time flows;
- 4) a new infrastructure based on DWDM will be established for long distance high-bandwidth (2,5 - 10 Gbps per wavelength) stream transport traffic patterns. This infrastructure will enable operators to set up different layer networks over different wavelengths. For example, given that a fibre offers $\lambda 1 - 80$, the operator can choose to use $\lambda 1$ as a signalling network, to use $\lambda 2$ as a traditional voice network, to use $\lambda 3 - 60$ for 58 separate high-quality SDH transport layer networks and finally use $\lambda 61 - 80$ as 20 separate best-effort GBE transport layer networks.

6.2 Problems with traffic management under packet switching

The question is to which extent the IP protocol will support messages, files and real-time flows under TCP or UDP. Considering that TCP was initially designed to serve the transmission of messages and files through a by definition unreliable network, it follows that TCP is fulfilling its role very well. TCP will guarantee content integrity, although it may take some time. UDP, however, will serve only those applications that are not concerned with proper delivery.

For real-time flows, however, the situation is different. TCP and UDP are not designed for real-time flows. TCPs retransmission functionality makes it unsuitable for such a task. UDP may do the work better concerning timing integrity, assuming that the network is not too congested, and reasonably reliable. The use of RSVP to perform bandwidth allocation for UDP traffic is the probable solution for real time flows in the IP environment.

At the time of this writing two new technologies are emerging aimed to solve the traffic issues, these are diffserv and MPLS.

Therefore, the conclusion is that given a transmission network that offers spare capacity enough to cope with traffic variations, the two protocol pairs IP and UDP/IP may reasonably well together fulfil the requirements over a transmission with overprovisioning in the order of 60 - 70 % and little overhead for traffic management.

6.3 Interworking cases; IP to fibre

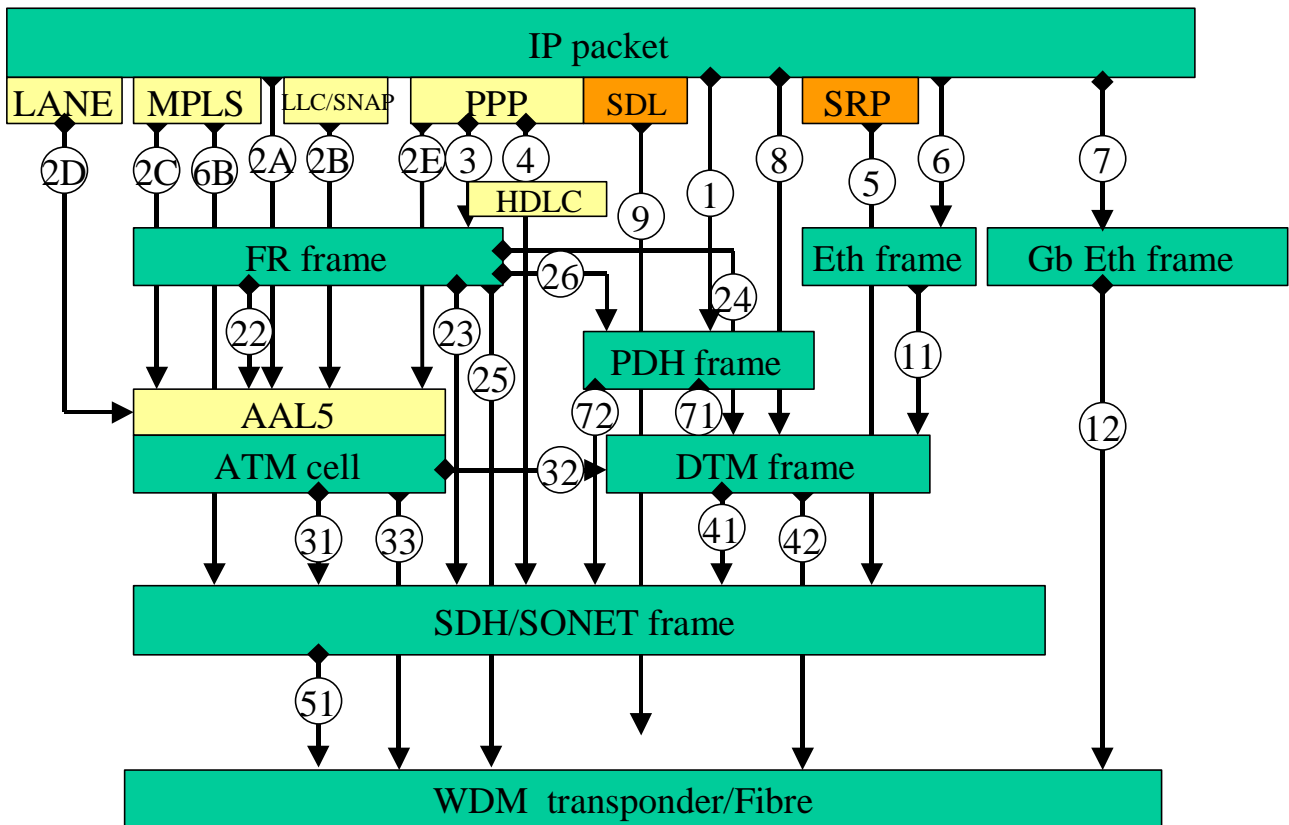


Figure 3: Different ways to transport IP directly or indirectly over Fibre

These are the cases according to figure 4:

Table 1

Case	Protocol data mapping	Standard reference
1	IP packet over PDH	
2A	IP packet over AAL5/ATM	ITU-T I.363.5 [4]. IETF RFC 1483 [5]. IETF RFC 2225 [6]. IETF RFC 1755 [7].
2B	IP packet over LLC/SNAP/AAL5/ATM	Same as above.
2C	IP packet over MPLS over AAL5/ATM	IETF RFC 2105 [23].
2D	IP packet over LANE over AAL5/ATM	ATM Forum af-lane-0084.000 [14]. ATM Forum af-lane-0112.000[15].
2E	IP packet over PPP over AAL5/ATM	IETF RFC 2364 [8].
3	IP packet over PPP over FR frame	IETF RFC 1973 [9].
4	IP packet over PPP over HDLC over SDH frame	IETF RFC 2615 [10].
5	IP packet over SRP over SDH frame	Proprietary solution.
6	IP packet over Ethernet frame	Essentially a LAN technique.
7	IP packet over Gigabit Ethernet frame	Emerging application.
8	IP packet over DTM frame	ETSI SPAN 13 work item. Emerging application.
9	IP packet over PPP over SDL over WDM transponder/fibre	Proposed to OIF.
11	Ethernet frame over DTM frame	ETSI SPAN 13 work item. Emerging application.
12	Gigabit Ethernet frame over WDM transponder/fibre	IEEE 802.3z [17] Fibre Channel physical interface.
22	FR frame over AAL5/ATM	FRF 5 and FRF 8 (see bibliography).
23	FR frame over SDH frame	In theory only.
24	FR frame over DTM frame	ETSI SPAN 13 work item. Emerging application.
25	FR frame over WDM transponder/fibre	Strange solution. In theory only.
26	FR frame over PDH frame	In theory only.
31'	ATM cell over SDH frame	ITU-T Recommendation G.707 [19]. ITU-T Recommendation I.432.1-2 [12], [13]. ETS 300 300 [11]. ATM Forum af-uni-0010.002 [24]. ATM Forum af-phy-0046.000 [27]. ATM Forum af-phy-0128.000 [25]. ATM Forum af-phy-0133.000 [26].
31''	ATM cell over SONET frame	ANSI T1.105-1995
32	ATM cell over DTM frame	ETSI SPAN 13 work item. Emerging application.
33	ATM cell over WDM transponder/fibre	ETSI ETS 300 299 [2]. ITU-T Recommendation I.432.1-2 [12], [13]. ATMF af-phy-0128.000 [25].
41	DTM frame over SDH frame	ETSI TM1 work item t.b.d. Emerging but temporary solution.
42	DTM frame over WDM transponder/fibre	ETSI TM1 work item t.b.d.
51	SDH/SONET frame over WDM transponder/fibre	SDH/SONET physical protocol.

6.4 IP over ATM

6.4.1 IP/AAL5/ ATM (Cases 2A, 2B, 2C, 2D, 2E)

ATM is based on switching by means of the establishment of logical connections. Logical connections in ATM are referred to as virtual channel connections (VCCs). It is the basic unit of switching in an ATM network. Bundles of VCCs can be organized into virtual path connections (VPCs), sharing the same end points. Thus, all of the cells flowing through all of the VCCs in a single VPC are switched along the same path. The concept of VPC enables simpler management of logical paths.

Currently ATM is mainly used as a supplement to IP networks in an effort to provide finer granularity access and reduce the amount of routing of individual IP packets of large files and in particular real-time flows. It remains to be seen if the principle of traffic type multiplexing over one physical medium will survive the competition of individual traffic type adjustment for distinct wavelengths over optical fibres.

This problem is a matter of traffic management and is well researched. The typical stack is IP/AAL5/ATM. The use of ABR is not showing any significant improvement over UBR, although there is much higher computational complexity. (FL97).

AAL3/4 is probably not economically feasible in comparison with AAL5. Too much overhead.

AAL2 is an emerging option, e.g. for mobile applications. See ITU-T Recommendation I.363.2 [28].

AAL1 seems to lose ground.

At present, one of the ways of trying to ensure a given QoS for an IP service is to guarantee a fixed bandwidth between pairs of IP routers for each customer (level 2 QoS management). ATM provides a way to do this with variable granularity (using the Permanent Virtual Channels (PVCs) set up using the ATM management system, or Switched Virtual Channels (SVCs) dynamically set-up, all within Virtual Paths (VPs)). ATM can also use statistical multiplexing to allow certain users to access extra bandwidth for short bursts. This can help to guarantee a fixed and arbitrary bandwidth from less than 1Mbit/s to several hundred Mbit/s to many different customers. In addition the fine granularity can enable IP routers to be connected into a logical mesh easily, thus minimizing delays from intermediate routers.

One other benefit in using ATM protocol is the possibility to have differentiated traffic contracts which offer various quality of service depending on the application requirements. For IP traffic which is by essence connection-less, the UBR (Unspecified Bit Rate) traffic contract is mainly used within ATM networks. Nevertheless, if IP applications require a particular QoS, especially for real-time constraints, it is possible to use other service categories such as CBR (Constant Bit Rate) or VBR-rt (Variable Bit Rate – real time) which are suitable for real-time applications.

IP over ATM can also be used as an implementation of MPLS or Tag switching. In this situation, the PVCs are not set up by intervention of the ATM management system, but dynamically by the Tag switching or MPLS protocol. The "Tag" is a number, assigned at an IP router at the edge of a Tag-switched or MPLS domain, which identifies a path across the network, so that the packets may be routed more quickly without having the routers to look-up the destination address in the IP packet. This tag can either be appended to the IP packet, or can be stored in the encapsulation frame when a suitable field exists. For MPLS or Tag-switching based on ATM, the Tag can be stored in the ATM VCI (RFC 2105 Tag Switching). The framing requirements are then similar to those for "classical IP over ATM".

There are many flavours of IP over ATM (e.g. classical IP over ATM, LAN emulation, Multiprotocol over ATM). Here we concentrate only on the more standardized and mature classical ATM over IP protocol (see [3], [4], [5], [6], [7]).

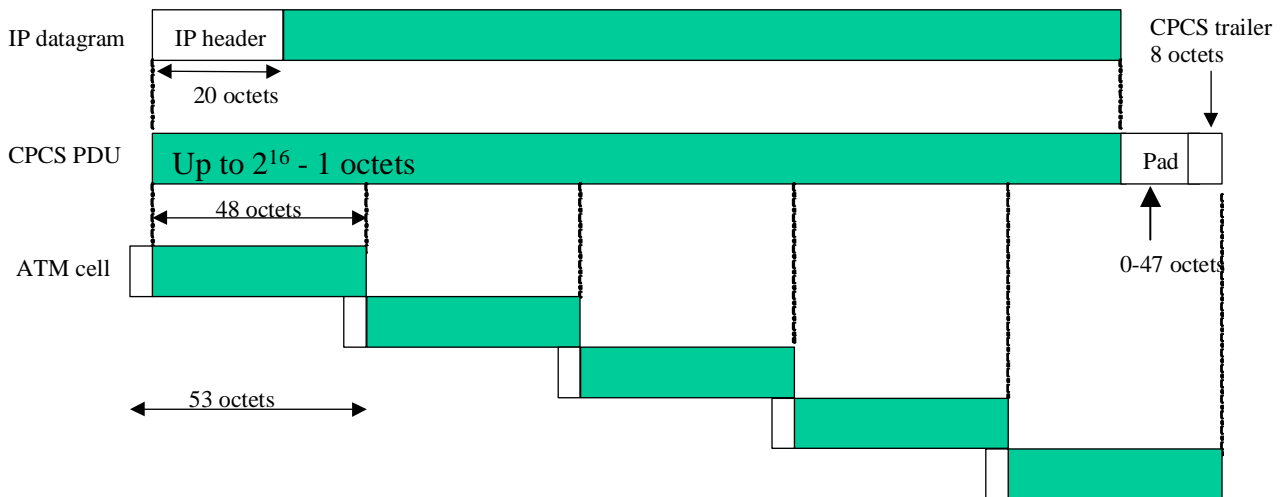


Figure 4: IP over AAL5/ATM

6.4.2 IP/LLC/SNAP/AAL5/ATM over SDH (Cases 2B + 31)

In this scenario, IP packets are segmented into ATM cells and assigned different Virtual Connections by the SDH/ATM line card in the IP router. The ATM cells are then packed into an SDH frame, which can be sent either to an ATM switch or directly to a WDM transponder for transport over the optical layer.

Table 2: Protocol stack for IP over LLC/SNAP over AAL5/ATM over SDH

IP	to be encapsulated, packets between 250 and 65 535 octets long. This packet includes already TCP or UDP header and IP header.
LLC/SNAP	Logical Link Control, [5]. Adds 8 bits overhead to IP packet to form ATM "PDU" (Protocol Data Unit) up to 65 535 octets long.
AAL5	Adds 8 overhead bits (length field, and 4-octets CRC) plus a 0 to 47 octets padding field, to form an AAL5 PDU, which fits into an integral number of 48-octets ATM payloads.
ATM	Segments AAL5 PDU into 48 octets payloads then adds 5-octets overhead to every 48-octets payload to form 53 octets ATM cells.
SDH	Puts ATM cells in SDH VC4 or concatenated VC4 payload [19]. Adds the SDH section overhead (81 bits including AU pointers), and a 9 VC4 bits Path overhead, to the 2 340 bits SDH VC4 payload. For concatenated VC4s, a V4-Xc payload is $X * 2\ 340$ long. ATM cells may cross VC4 boundaries, and their payload is scrambled with a $1 + x^{43}$ polynomial to provide sufficient transition density to allow for SDH clock recovery. The usual x^7 scrambling is used on top for the SDH payload. The SDH section and path overheads contain identifiers and error checking fields (BIP-n) for performance monitoring, as well as communication channels for managing the transport network.

See [5], [19].

Mapping of ATM over SONET/SDH physical layer is defined in several standardization bodies for various bit rates (bit rates below 155 Mbit/s are not mentioned here):

ETS 300 300 [11] defines ATM over SDH interfaces for 155 and 622 Mbit/s operation at S_B and T_B reference points.

ITU-T Recommendation I.432.1-2 [12], [13] defines ATM over SDH interfaces for 155 and 622 Mbit/s operation at the S_B and T_B reference points.

ITU-T Recommendation G.707 [19] provides the general mapping of data tributaries into VC4-nc virtual containers enabling the definition of ATM transport on SDH physical layers at and above 155 Mbit/s for NNI interfaces.

ATM Forum AF-UNI-0010.002 [24]: UNI 3.1 which defines ATM transport at 155 Mbit/s over SONET/SDH both at UNI and NNI interfaces,

ATM Forum AF-PHY-0046.000 [27] defines ATM transport at 622 Mbit/s over SONET/SDH both at UNI and NNI interfaces.

ATM Forum AF-PHY-0128.000 [25] is under development and defines ATM transport at 2 488 Mbit/s over SONET/SDH both at UNI and NNI interfaces.

6.4.3 IP/LLC/SNAP/AAL5/ATM over fibre (Cases 2B + 33)

This scenario is identical to the previous one, from an architectural point of view. The only difference is that ATM cells are not encapsulated into SDH frames, instead there are sent directly on the physical medium by using an ATM cell-based physical layer.

Cell-based physical layer is a relatively new technique (compared to SONET/SDH) for ATM transport. Cell-based physical mechanisms have been developed specifically to carry the ATM protocol; this technique can not support any other protocol except if these protocols are emulated over ATM.

ATM cell-based physical layers are defined in several standardization bodies for various bit rates (bit rates below 155 Mbit/s are not mentioned here).

ETSI ETS 300 299 [2] defines ATM cell-based physical interfaces for 155 and 622 Mbit/s operation at S_B and T_B reference points.

ITU-T Recommendation I.432.1-2 [12],[13] also defines ATM cell-based physical interfaces for 155 and 622 Mbit/s operation at the S_B and T_B reference points.

ATM Forum AF-PHY-0128.000 [25] is under development and defines ATM transport at 622 and 2 488 Mbit/s over cell-based physical layers, both at UNI and NNI interfaces.

Cell-based physical layer is dedicated to ATM transport. In the transmitter, the physical layer receives ATM cells coming from the ATM layer and inserts idle cells when there is no ATM layer cell to transmit. Then all the cells are scrambled before being sent on the physical medium. Cell-based physical layer uses the DSS (Distributed Sample Scramble) which scrambles both the cell payload and the cell header (except the HEC field); this insures robustness against false cell delineation and good transmission performance. In the receiver cell delineation is achieved by checking continuously the HEC syndrome in the received bit stream. Physical layer F1 and F3 cells are dedicated to cell-based physical layer for operation and maintenance of the transmission path and the regenerator section.

The cell-based physical layer offers the following functions:

- indication of AIS (Alarm Indication Signal) and RDI (Remote Defect Indication) both at the path and regenerator levels;
- BIP8 calculation over consecutive blocks of cells to detect any transmission error;
- possible support of an in-band communication channel (through reserved bits in F1 and F3 OAM payloads);
- indication to the far end of the number of errored blocks detected in one direction of transmission. This function enables both ends to have the same view of the performance measured on the connection;
- 1 + 1 protection mechanism at the physical layer, without disruption. This mechanism is under specification in the ITU-T.

6.4.4 IP/LANE/AAL5/ATM (Case 2D)

This case puts IP on an intermediate layer (LANE) which is carried over AAL5 and ATM (see [14], [15]).

6.4.5 IP/PPP/AAL5/ATM (Case 2E)

See [8], [20].

Table 3: Evaluation summary of IP over ATM

Connection principle	IP: Connectionless. ATM: Virtual path and virtual channel. Mainly statistical multiplexing, but reserved channel can be established.
Bandwidth flexibility	ATM is in practice always using permanent virtual channels. Bandwidth of a permanent channel cannot be changed by the user, but active ATM cells can be filling the channel up to allocated bandwidth. The choice between CBR and VBR allows some flexibility. Applications can share an existing channel.
Overhead economy	IP header + padding + CPCS trailer + cell headers. Approximately 20 overhead.
Content integrity	Guaranteed for reserved channels. Under statistical multiplexing cannot be guaranteed.
Timing integrity	Variable buffer delays may destroy timing integrity.
Multicast support	None. Always point-to-point connections in ATM flows.
Major advantage	Better granularity than IP directly over PDH or SDH/SONET. Higher speed than Frame Relay.
Major problem	Too much processing of ATM cells. Non-determinism . Traffic management complicated.

6.5 IP over PDH (Case 1)

The transport of IP over PDH can be done using adaptation provided by either PPP, ATM or proprietary solutions.

6.6 IP/PPP over Frame Relay (Case 3)

Frame Relay is a high-speed packet switching technology which has evolved from the development of narrow-band ISDN. The maximum connection address speed is 2 Mbps, as in primary rate ISDN. The access protocol is based on LAPD. Multiplexing of logical channels is a layer 2 function and there is no need for layer 3 processing. Compared to its predecessor, X.25, only of subset of layer 2 functions are defined. The remaining functions consists of frame delimiting and transparency, error detection and virtual circuit multiplexing. Other layer 2 functions of X.25 such as error recovery, flow control and timer maintenance are left for the termination nodes of the frame relay connection, and thus not part of the frame relay network.

Frame relay networks are connection-oriented and use out-of-band signalling. This feature makes Frame Relay a traditional telecommunications protocol, where the availability of a common channel signalling system often is taken for granted.

A Frame relay connection can neither guarantee content integrity nor timing integrity. Essentially, it is a best-effort approach. Concerning the suitability for various traffic types, short frames can be used for real-time flows, since it does not matter so much if some frames are lost. In order to serve file transfer well, there is a need for a windowing mechanism in the terminal nodes. This function can be supplied by TCP over IP. However, if TCP/IP are used under a user application, it is not likely that IP packets will be mapped onto FR frames. Thus, IP over Frame Relay is a circumvention case that cannot be justified in practice(see [9]).

Table 4: Evaluation summary

Connection principle	Connectionless.
Bandwidth flexibility	No flexibility.
Overhead economy	Depending on frame length. Each frame demands only 6 to 8 octets. Therefore good overhead economy.
Content integrity	Under statistical multiplexing cannot be guaranteed.
Timing integrity	Relay delays may vary. No guarantee. Under 50 % load good throughput.
Multicast support	No support.
Major advantage	Simple transmission technique.
Major problem	Low bandwidth. Non-determinism Out-band signalling needed.

6.7 IP over DTM (Case 8)

DTM is based on fast circuit switching over a shared medium, normally a dual fibre ring. There is no problem with traffic management for large files and in particular real-time flows. Thus DTM is in this sense completely complementary to IP, and a competitor to ATM. The problematic traffic types for DTM are messages and small files, where the establishment of a dedicated circuit of 512 kbps cannot be justified. The solution is a hybrid of routing and switching. DTM has a solution for IP over DTM which offers an edge routing approach to messages and small files.

It can thus be said that when IP needs a switching based solution to handle large files and real-time flows efficiently, then DTM needs a routing based solution to handle messages and small files efficiently.

DTM is allocating necessary slots of 512 kbps each for a channel for the transport of user data. Each channel is a separate entity for which there is a guarantee for content and timing integrity.

As concerns multicasting, DTM is a system based on a shared medium, since the frames are passing sequentially the connected access nodes or switches. It has thus a scaling advantage concerning multicasting, under the restriction that the full benefit is gained only if there is DTM end-to-end. The volume of traffic grows only linearly with the number of engaged logical connections (e.g. rings) regardless of how many nodes are receivers in every single connection.

Table 5: Evaluation summary

Connection principle	Circuit-switched. Permanent channels can be used for routed packet transfer best-effort.
Bandwidth flexibility	From 512 kbps up to port bandwidth in increments of 512 kbps.
Overhead economy	1 - 2 % of bandwidth used for signalling.
Content integrity	Guaranteed.
Timing integrity	Guaranteed.
Multicast support	Supported.
Major advantage	Simple time-division multiplexing transmission technique.
Major problem	Messages cannot afford their own channels. Have to be transmitted as IP packets over permanent DTM channels.

6.8 Frame Relay over ATM (Case 22)

The Network Interworking Function (IWF) facilitates the transparent transport of frame relay user traffic and frame relay PVC signalling traffic over ATM, connecting two frame relay networks over an ATM backbone. This means that multiprotocol encapsulation and other higher layer procedures are transported transparently as they would over leased lines.

Network Interworking uses an ATM network to connect two frame relay networks. The Network IWF can be external to the networks but is more likely to be integrated into the ATM network switch or frame relay switch. Each frame relay PVC can be carried over an ATM PVC, or all of the frame relay PVCs can be multiplexed onto a single ATM PVC.

The Service Interworking Function

The Service IWF does not transport traffic transparently, but instead functions more like a protocol converter facilitating communication between dissimilar frame relay and ATM equipment. A frame relay user sends traffic on a PVC through the frame relay network to the Service IWF which then maps it to an ATM PVC. The frame relay PVC address-to-ATM PVC address mapping and other serving options are configured by the network management system associated with the IWF. The Service IWF can be external to the networks, but is more likely to be integrated into the ATM network switch or frame relay switch. Note that in the case of Service Interworking, there is always one ATM PVC per frame relay PVC.

Due to architectural differences between frame relay and ATM, the Service IWF necessarily converts header function mapping and multiprotocol data unit headers from frame relay formats to ATM formats and vice versa (see bibliography).

Table 6: Evaluation summary

Connection principle	Permanent virtual circuit.
Bandwidth flexibility	None.
Overhead economy	ATM overhead adds to FR overhead.
Content integrity	No guarantee.
Timing integrity	No guarantee.
Multicast support	None.
Major advantage	Common ATM backbone used, will eliminate dedicated FR network.
Major problem	None.

6.9 Ethernet over DTM (Case 11)

This case is implemented as DTM LAN Emulation (DLE) according to VLAN standard (IEEE 802.1Q [29]). It needs a particular DLE server for address resolution.

IEEE 802 Local Area Networks (LANs) of all types may be connected together with Media Access Control (MAC) Bridges, as specified in ISO/IEC 15802-3 [30]. It defines the operation of Virtual LAN (VLAN) Bridges that permit the definition, operation and administration of Virtual LAN topologies within a Bridged LAN infrastructure.

Different users may be kept separate through the DTM network from Ethernet switch to Ethernet switch. DTM layer network will serve as carrier supporting VLAN.

Table 7

Connection principle	Circuit-switched.
Bandwidth flexibility	From 512 kbps up to port bandwidth in increments of 512 kbps.
Overhead economy	1 - 2 % of bandwidth used for signalling.
Content integrity	Guaranteed.
Timing integrity	Guaranteed.
Multicast support	Supported.
Major advantage	Simple time-division multiplexing transmission technique.
Major problem	None.

7 Stream transport in an integrated transport network

7.1 Basic assumptions

The case for stream transport is actual when the traffic management functions have formed the payload into a large multiplexed stream for which the service requirements shall fulfil the most demanding of the component traffic real-time flows. E.g. if the multiplexed stream contains real-time flows, the real-time requirements must be fulfilled concerning delay. If the multiplexed stream contains files, then no bits must be lost. If there is no control over which traffic types are involved in a multiplexed stream, then it is obvious that the stream must fulfil the union of all possible service requirements. This means that a generic transport network must offer a synchronous transmission service free of bit loss.

This strong requirement is today's reality. However, since in theory it should be possible to let some wavelengths in a WDM system carry multiplexed real-time flows and other carry only multiplexed files, it should be possible to define relaxed requirements for distinct wavelengths.

Under the assumption that it is possible for the application to mark the service requirements for a traffic type instance (compare IETF DiffServ) it should also be possible to consider this marking through the traffic management layer and engage wavelengths according to particular service classes, that could match the traffic types as follows:

- 1) extremely reliable permanent low-bandwidth connection for Message;
- 2) asynchronous extremely reliable high-bandwidth connection for File;
- 3) strongly synchronous high-bandwidth connection for Real-time flow.

Considering the special round-trip delay requirements for voice and interactive video, it should be possible to allocate special connection bundles for all channels (traffic instances) required for distinct video conference sessions. Then these channels should be open also for idle packets that are sent for silence. On the other hand, there is a certain cost for such a luxury.

There are several options available for stream transport of IP packets over a physical medium. In those cases, the IP packets are multiplexed into a stream and thus not available as individual flows. The stream transport is based on the assumption that capacity is available in the fibre or the virtual container. If capacity is temporarily not available, then the payload is discarded. Note that this situation also holds for IP payload that has been intermediately managed and shaped through ATM, Frame Relay or PDH.

The major stream transmission protocols are SDH/SONET and Gigabit Ethernet. SDH/SONET as a synchronous technique is typically a telecommunications based protocol. Gigabit Ethernet is a typical datacom based protocol and a fully compatible extension to Ethernet and Fast Ethernet.

7.2 IP over SDH

7.2.1 IP/PPP/HDLC over SDH or SONET (Case 4)

This is an approach favoured by many operators in a situation where there is no need for traffic management. It is in USA called "Packet over SONET".

It is possible to simply use SDH formats to frame encapsulated IP packets for transmission over WDM, probably using a transponder (wavelength adapter), or it is also possible to transport the SDH-framed IP over an SDH transport network along with other traffic, which may then use WDM connections.

SDH can currently be used to protect IP traffic connections against cable breaks by automatic protection switching (APS) in a variety of guises. This can also be performed in the optical layer, although only 1 + 1 OMSP or OChP are currently commercially available.

The line card in the IP router performs the PPP/HDLC framing. The optical signal is then suitable for short transmission over optical fibre either into an SDH network element, a neighbouring IP router, or a WDM transponder for further transmission. There are also different types of IP over SDH interfaces:

- VC4 or Concatenated VC4 "fat pipes" which provide aggregate bandwidth without any partitions between different IP services which may exist within the individual packet flows;
- Channellized interfaces, where an STM16 optical output may contain 16 individual VC4s, with separate IP bandwidth in each;

The version of IP over SDH examined here uses PPP encapsulation and HDLC framing. The protocols used are shown in table 8 (see [10]).

Table: 8 Protocols used for IP over SDH

IP	Datagram with maximum length of 65 535 octets.
PPP	PPP encapsulation, RFC 1661 [16]. Adds 1 or 2 octets "protocol field" and optional padding. PPP also provides for a link establishment protocol, which is not a critical function for IP over SDH.
HDLC	Framing, RFC 1662 [31]. Adds a flag bite indicating frame start, 2 more overhead bits, and a 2 bits frame check sequence (FCS), resulting in a frame up to 1 500 octets long. Together with PPP, HDLC results in a 7 or 8 octets overhead being added to the IP packet.
SDH	Put HDLC frames into a VC4 or concatenated VC4 payload, RFC 2615 [10]. Adds the SDH section overhead (81 bits including AU pointers), and a 9 VC4 bits Path overhead, to the 2 340 bits SDH VC4 payload. For concatenated VC4s, a V4 - Xc payload is X* 2 340 long. The frames are allowed to cross VC4 boundaries. An internet draft [PPP_SONET] specifies using a scrambling $1 + x^{43}$ polynomial, like ATM, to minimize the risk that a malicious user may send data which may cause SDH to loose synchronization.

Table 9: Evaluation summary

Connection principle	Circuit-switched. Add-drop of tributaries possible.
Bandwidth flexibility	According to tributary hierarchies. From 2 Mbps and upwards to 2,5 Gbps virtual containers.
Overhead economy	About 6 % of bandwidth used for overhead in containers.
Content integrity	Guaranteed.
Timing integrity	Guaranteed.
Multicast support	Not supported.
Major advantage	Simple time-division multiplexing transmission technique. Very reliable and proven technique.
Major problem	Network management is not standardized. This tends to drive towards one-supplier solutions.

7.3 IP over SRP over SDH (Case 5)

Spatial Reuse Protocol (SRP) [18] is a system for transmission of IP packets over fibre. Its basic transmission principle is synchronous transmission of data in (normally dual) fibre rings. Its traffic policing principle is queuing and statistical multiplexing. The central protocol is SRP that means that transmission capacity is used only when needed from end-to-end. The only exception is multicasting, where the capacity is managed by the sending node.

SRP has protocols for ring resilience, which allow for reconfiguration at link failure in less than 50 ms.

The SRP system is based on a dual ring topology. A ring consists of two fibres, where the traffic flow is going in one direction on one fibre and in the opposite direction on the other fibre. Both rings can transport data and control packets. The signalling for one traffic stream is going in the opposite direction of the data stream.

SRP utilizes SONET/SDH framing. This transport solution enables SRP to interwork with SONET/SDH equipment for transport of data, but also to use the SONET/SDH framing specification without using the other aspects of SONET/SDH. Since SONET/SDH is synchronous, this means that although the packets are entered asynchronously into the system in queues, the transport of data is synchronous.

Thus SRP is one of the possible protocol layers over SONET/SDH. IP services can be initially deployed using spare capacity on SONET/SDH equipment. But also, as the IP traffic grows it can be migrated to packet optimized transport based either on dark fibre or WDM equipment. The physical layer interface, which is then used, is not published.

SRP rings spanning long distances may use regenerators (this is an optical system feature, and not within the scope of SRP analysis).

A central feature of SRP is the Spatial Reuse Protocol (SRP). It derives its name from the protocol's destination-stripping operation. SRP utilizes **destination stripping**, which means that the destination node strips the packets from the ring and the full bandwidth on other segments of the ring is available for utilization by other packets.

SRP is a best-effort approach to traffic shaping. SRP performs the traffic shaping by means of two priority queues at the access node. Highest priority queue is instructed from higher layer protocols to put in higher priority packets according to some kind of service level agreement or Class of Service coding. Lower priority queue will take the rest of the traffic. Both queues will release data to the fibre according to statistical multiplexing under a certain level of overprovisioning. It is uncertain whether reservations schemes are used or feedback mechanisms we know from ATM such as Available Bit Rate is used.

The major principle behind the priority and traffic management is to utilize the raw fibre capacity and assume that the available bandwidth always is enough. Within one ring and if all sources are under control, this may be a very effective principle.

Each node gets its fair share of ring bandwidth by controlling the rates at which packets are forwarded onto the ring from upstream sources versus the rate at which packets are sourced by the ring node.

SRP uses statistical over-subscription with highly elastic burst handling capabilities. Service providers also have the flexibility to engineer over-subscription factors to fit the ring application (such as access rings versus backbone rings).

WDM introduces a new layer between the transmission system and the physical fibre. The fact that several logical fibre channels are multiplexed into one physical fibre using a certain number of distinct wavelengths (now 32 or 40).

The SRP access node is considering each wavelength in a WDM system as one logical and also physical link. Thus the multiplexing of a number of wavelengths into one fibre by the WDM equipment is not part of the SRP system. Each wavelength offered as a link by the WDM system will be physically connected to the access devices within the SRP system.

Table 10: Evaluation summary

Connection principle	Best effort over two level priority queues.
Bandwidth flexibility	Flexible.
Overhead economy	Similar to SDH/SONET.
Content integrity	No guarantee. Best effort. Overload will lead to discarding.
Timing integrity	No guarantee.
Multicast support	Supported.
Major advantage	Simple technique for two-level priority synchronous transmission near fibre.
Major problem	Proprietary. Not a network layer solution.

7.4 IP over Gigabit Ethernet over WDM transponder/fibre (Cases 7 + 12)

Gigabit Ethernet uses the same transmission schemes and frame formats as Ethernet (10 Mbps) and Fast Ethernet (100 Mbps). It provides a MAC layer that supports IP. It has no bandwidth flexibility and is mainly used for backbones between edge or access nodes supporting Fast Ethernet.

Ethernet accounts for over 85 % of LANs world-wide, and the new Gigabit Ethernet standard can be used to extend high-capacity LANs to MANs and maybe even WANs, using Gigabit line cards on IP routers, which can cost 5 times less than SDH line cards with similar capacity. For this reason, Gigabit Ethernet could be a very attractive means to transport IP over "metropolitan" WDM rings, for example, or even over longer WDM connections. Furthermore, 10 Gbit/s Ethernet ports are likely to be standardized in the near future. Lower-bit rate Ethernet networks (e.g. 10Base-T or 100Base-T) have been used a lot in a half-duplex mode, where the bandwidth available for transmission is shared between all users, and between both directions of transmission. To police access to the shared bandwidth, CSMA-CD is used (see IEEE 802.3z [17]). This imposes limits on the physical size of the network, where the transit time cannot exceed the "slot time" which is the minimum frame length (512 bits for 10Base-T and 100Base-T). For a bit rate of 1Gb/s, using a minimum frame length of 512 bits would imply an Ethernet network only roughly 10m long, and for this reason the minimum frame length has been redefined to be 4 096 bits for Gigabit Ethernet. However, this still limits the network size to 100m, so full duplex mode is more attractive using Gigabit Ethernet. When Gigabit Ethernet (1000Base-X) is used in full-duplex mode, it becomes simply an encapsulation and framing method for IP packets, and the CSMA/CD functionality is not used. Ethernet switches can also be used to extend the network topology beyond a point-to-point connection.

The frame structure for Gigabit Ethernet is shown in figure 5. The maximum payload length can be extended to 9 000 octets with Gigabit Ethernet (Jumbo Frames).

Idle	12
Preamble	7
Start frame delimiter	1
Destination address	6
Source address	6
Frame length	2
Logic link control field + payload (max length 1500 octets)	
Frame check sequence	4

Total overheads: 38

Figure 5: Gigabit Ethernet Frame

The Ethernet frame is encoded onto an optical carrier using 8b/10b encoding. 10 bits are used to represent every octet. For a 1Gb/s throughput, the line rate is therefore 1,25Gb/s. The 8b/10b encoding is designed to always present a high density of transitions in the optical signal, so that clock recovery is easier. When packets are not being transmitted, the idle period is filled with symbols having a sufficient density of transitions between 1 and 0 logical states to ensure that clock recovery is maintained (see [17]).

Table 11: Evaluation summary

Connection principle	Packet transfer point-to-point.
Bandwidth flexibility	Full flexibility for packets in the sense that they all enter a queue and then are transmitted in the common medium. In the medium they cannot be related to user applications.
Overhead economy	11 % overhead.
Content integrity	No guarantee. Best effort. Overload will lead to discarding.
Timing integrity	No guarantee.
Multicast support	Not supported. This is a backbone technique.
Major advantage	Simple one-queue-based transmission technique.
Major problem	Any multiplexed traffic type demanding high QoS may suffer.

7.5 IP over PPP over SDL over fibre (Case 9)

Simple Data Link (SDL) is a framing method proposed, and can replace HDLC framing for PPP-encapsulated packets. The SDL format can then be inserted into an SDH payload for transmission over WDM or SDH equipment. The SDL format can also be encoded directly onto an optical carrier: SDL specifies the bare minimum functionality to be able to do this.

SDL has been proposed for international standardization both in the IETF (Internet Engineering Task Force) and in the OIF (Optical Interworking Forum).

The information provided in this clause is based on RFC 2615 [10]).

Transmitter operation

SDL interface is packet oriented; each SDL packet actually contains a PPP packet.

Each PPP packet is preceded by a header field which indicates the packet length and a CRC field which protects the header field. Each PPP packet is followed by a CRC field which protects the PPP packet itself. The interface structure is shown below:

- packet length -> 2 octets;
- header CRC -> 2 octets;
- PPP packet -> n octets;
- packet CRC -> 4 octets.

This SDL packet structure is repeated continuously by the transmitter.

Two actions are performed by the receiver:

- recover SDL packet boundaries;
- descramble PPP packets.

Synchronization process

SDL operation mode can be enabled by two methods: an LCP (Link Control Protocol) method and a prior-arrangement method. For these two methods the PSL (Path Signal Label) field which is bite C2 of the SONET/SDH overhead is used to indicate the type of the data carried on the path. The first method to set up SDL allows easy configuration and compatibility with existing equipment. The second method allows general use with separate SONET/SDH transmission equipment with PSL limitations.

Table 12: Evaluation summary

Connection principle	Packet oriented point-to-point.
Bandwidth flexibility	Full flexibility.
Overhead economy	The SDL overhead amounts to 3 % based on an average IP packet length of 350 bits.
Content integrity	Guaranteed. CRC check on both PPP header and packet.
Timing integrity	Not guaranteed.
Multicast support	Not supported.
Major advantage	SDL is a low complexity protocol with a reduced overhead as compared to HDLC and SDH.
Major problem	No monitoring of the optical channel. No protection switching protocol.

8 Comparison of available bandwidth

One way to compare framing methods is to calculate the overhead. On the basis of an "average length" 350 bits long IP packet is calculated the overhead of the mentioned encapsulation methods and the results are summarized in table 13.

Table 13

Encapsulation/ Framing	Overhead	Connection capacity Mb/s
IP/ATM/SDH	28 %	1 944
IP/ATM/fibre	23 %	2 016
IP/PPP/SDH	6 %	2 335
IP/SDL	3 %	2 411
IP/GbE	11 %	902
IP/DTM	t.b.d.	1 000

Summary of overhead percentage and connection capacity for different encapsulation methods. The connection is STM-16 for the first 3 methods, and GbE and DTM with a line rate of 1,25Gb/s for the two last.

Annex A (informative): Bibliography

- IETF RFC 1889 (1996): "A Transport Protocol for Real-Time Applications".
- ITU-T Recommendation Q.922 (1992): "ISDN data link layer specification for frame mode bearer services".
- FL97
C. Fang and A. Lin "TCP performance in ATM Networks: ABR parameter tuning and ABR/UBR comparisons".
Proceeding IEEE Singapore International Conference on Networks, April 1997.
- GIGA 98
- N. Giroux and Sudhakar Ganti "Quality of Service in ATM networks".
- Prentice-Hall 1998
- William Stallings "High-speed networks: TCP/IP and ATM design principles", Prentice Hall, 1998.
- M. Sexton and A. Reid "Broadband networking: ATM, SDH and SONET", Artch House, 1997.
- Andrew S. Tanenbaum "Computer Networks". Third edition, Prentice Hall, 1996.
- Net Insight " Introduction to DTM", White paper 1998. Available at www.netinsight.se.
- FRF.5 Frame Relay/ATM Network Interworking Implementation - December 1994.
- FRF.8 Frame Relay / ATM PVC Service Interworking Implementation Agreement - April 1995.

NOTE: The following was received on May 6, 2000 from: Albert Herrera.
Cisco's IPR disclosure statement pertaining to the Spatial Reuse Protocol (SRP). Cisco has a patent pending that may relate to the work underway or work derived from draft-tsiang-srp-02.txt. Upon approval by the IESG of the relevant Internet standards track specification and if any patents issue to Cisco or its subsidiaries with claims that are necessary for practicing this standard, any party will be able to obtain the right to implement, use and distribute the technology or works when implementing, using or distributing technology based upon the specific specification(s) under openly specified, reasonable, non-discriminatory terms.
Requests may be sent to:
Robert Barr
300 East Tasman Dr.
San Jose, CA 95134 USA
phone: +1 (408) 525 9706

History

Document history		
V1.1.1	February 2001	Membership Approval Procedure MV 20010406: 2001-02-06 to 2001-04-06
V1.1.1	April 2001	Publication