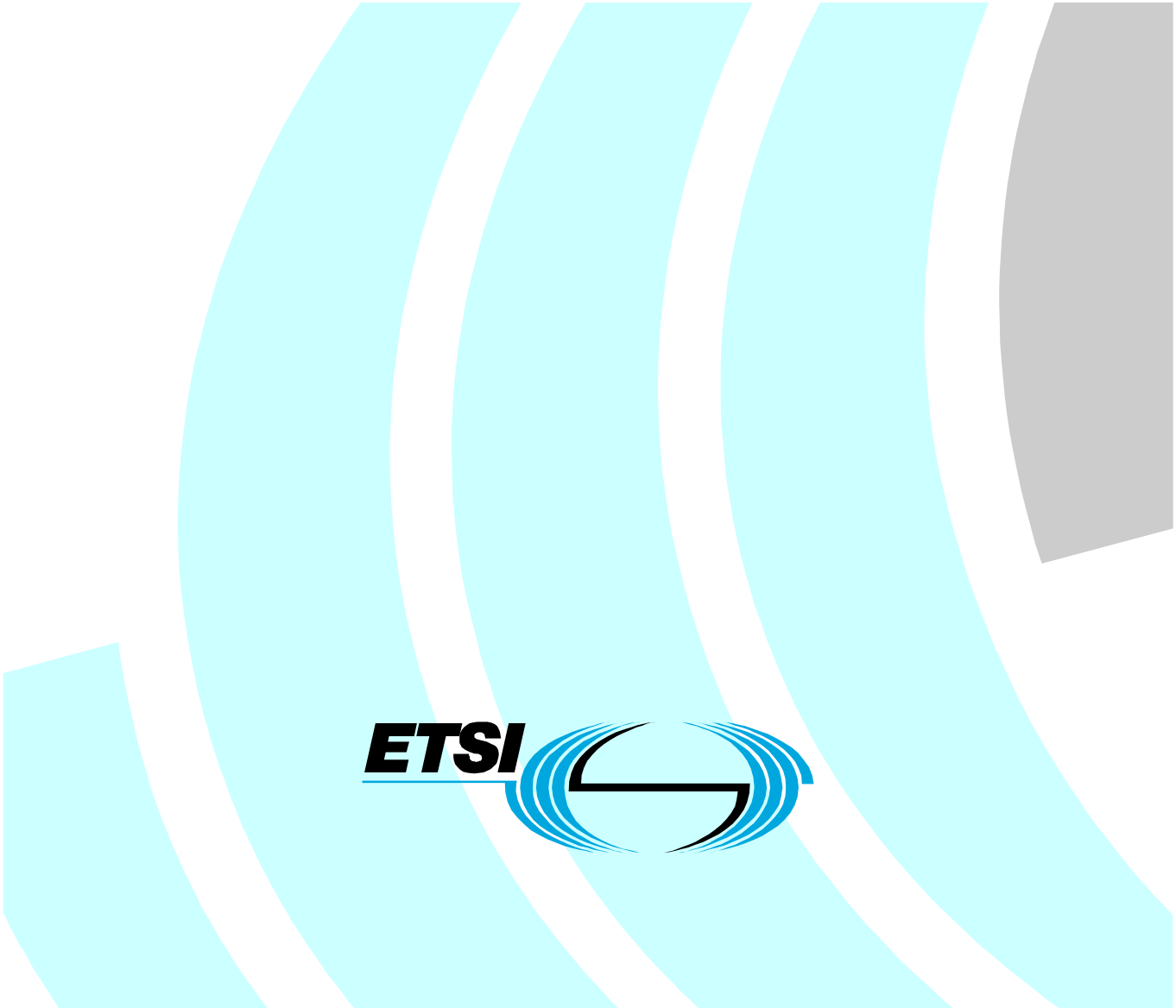
ETSI TS 102 462 V1.1.1 (2006-12)

Technical Specification

Satellite Earth Stations and Systems (SES); Broadband Satellite Multimedia (BSM); QoS Functional Architecture



Reference
DTS/SES-00102

Keywords
broadband, interworking, IP, multimedia,
QoS, satellite

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

Contents

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://webapp.etsi.org/IPR/home.asp).

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 Technical Specification (TS) has been produced by ETSI Technical Committee Satellite Earth Stations and Systems (SES).

Introduction

IP-based services and their users are growing ever more sophisticated, and QoS is a feature which will be increasingly valuable for service differentiation and support of more sensitive applications (e.g. MMoIP). In contrast to wired or optical networks where over-provisioning of capacity is often used to ensure QoS for packet-based transport, in satellite systems, as in other wireless networks and access networks in general, capacity is precious and must be carefully allocated. This requires more sophisticated QoS methods which are closely linked to resource provision and control at lower protocol layers than IP.

QoS provision within BSM systems themselves is one of the first aims, but since BSM systems are intended for use within the Internet, end-to-end QoS across integrated networks including satellites is also a key aspect that will be needed to be enabled by BSM systems.

The general issues concerning Quality of Service (QoS) and architectures in BSM systems are described in TR 101 984 (see Bibliography). TR 101 985 (see Bibliography) describes further specific QoS requirements. TR 102 157 (see Bibliography) describes functional models for QoS concerning IP-over-satellite aspects.

The present document takes the requirements of the above documents and defines functional architectures that are needed within BSM systems to implement end-to-end QoS for IP-based applications with the potential to run over integrated network scenarios. Compatibility with QoS requirements for generic internetworking, such as those from IETF, ETSI TISPAN (see Bibliography) are taken into account .

The BSM architecture is characterised by the separation between common Satellite-Independent (SI) protocol layers and alternative lower Satellite-Dependent (SD) layers [3]. At the SI layers, several methods of ensuring end-to-end QoS over integrated networks are foreseen, by means of signalling protocols (e.g. based on SIP (RFC 3261, see Bibliography), NSIS (see Bibliography), etc.) t the session (or application) layers and DiffServ, RSVP/IntServ at the IP layer. At the SD Layers alternative lower protocol layers offer different QoS characteristics. The focus of the architecture definition here is on maintaining compatibility with these alternative methods and approaches by addressing the generic BSM QoS functions required in the IP and Satellite-Independent layers. These functions will provide interfaces where appropriate with higher-layer and lower-layer QoS functions, and with external networks and customer equipment.

1 Scope

The present document defines a BSM functional architecture required to provide multimedia Quality of Service to the end user. This architecture identifies the functional elements to allow QoS provision in BSM systems integrated with heterogeneous networks. It includes the interfaces of these elements to other QoS functional elements in adjacent networks and customer equipment, and in higher or lower protocol layers.

The multimedia services targeted are based on the Internet Protocol (IP) and use the QoS capabilities of IP as far as necessary, without relying on the support of other underlying protocols (such as MPLS). QoS for unicast rather than multicast services is the primary focus.

A key feature of the BSM architecture is the SISAP. The way in which the QoS functions interact across the SISAP is also illustrated in the architecture.

The approach adopted is to define the following aspects in sequence:

- the QoS service requirements;
- the global QoS functional and network architecture requirements in which BSM systems will play a role;
- the detailed BSM QoS architecture of main functional blocks (ST's, etc.).

Several architectural cases are illustrated for QoS provision. These cases are intended to show the potential evolution from a simple QoS solution with quasi-static resource allocations to more sophisticated services with dynamic resource reservation.

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.

Referenced documents which are not found to be publicly available in the expected location might be found at http://docbox.etsi.org/Reference.

NOTE: While any hyperlinks included in this clause were valid at the time of publication ETSI cannot guarantee their long term validity.

[1] ETSI TS 102 292: "Satellite Earth Stations and Systems (SES); Broadband Satellite Multimedia (BSM) services and architectures; Functional architecture for IP interworking with BSM networks".

[2] IETF RFC 2475: "An Architecture for Differentiated Services".

NOTE: Available at http://www.ietf.org/rfc/rfc2475.txt

[3] ETSI TS 102 357: "Satellite Earth Stations and Systems (SES); Broadband Satellite Multimedia (BSM); Common Air interface specification; Satellite Independent Service Access Point SI-SAP".

[4] IETF RFC 2474: "Definition of the DS Field".

NOTE: Available at http://www.ietf.org/rfc/rfc2474.txt

[5] IETF RFC 2597: "Assured Forwarding PHB Group".

NOTE: Available at http://www.ietf.org/rfc/rfc2597.txt

[6] IETF RFC 2598: "An Expedited Forwarding PHB Group".

NOTE: Available at: http://www.ietf.org/rfc/rfc2598.txt

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the terms and definitions given in ETSI, IETF, ITU, and the following apply:

architecture: abstract representation of a communications system.

NOTE: Three complementary types of architecture are defined:

- **Functional Architecture**: the discrete functional elements of the system and the associated logical interfaces.
- **Network Arhitecture**: the discrete physical (network) elements of the system and the associated physical interfaces.
- **Protocol Architecture:** the protocol stacks involved in the operation of the system and the associated peering relationships.

bearer service: type of telecommunication service that provides the capability for the transmission of signals between user-network interfaces

behaviour aggregate: collection of packets with the same DS code point crossing a link in a particular direction

BSM bearer service: telecommunication service that a BSM subnetwork provides between a pair of SI-SAP's in different ST's

Best-Effort (BE) service: service that offers no QoS guarantees, just end-to-end connectivity

NOTE: When using queuing to prevent congestion BE queues are always the first ones to experience packet drop.

Class Of Service (COS): defines a way to divide traffic into separate categories (classes) to provide (e.g. Diffserv) to each class within the network

classification: examination of a packet to determine the CoS to which the packet should belong

code point: specific value of the DSCP portion of the DS field

NOTE: Recommended code points should map to specific standardised PHBs. Multiple code points may map to the same PHB.

connection oriented: communication method in which communication proceeds through three well-defined phases:

- connection establishment;
- data transfer; and
- connection release.

connectionless: communication method that allows the transfer of information between users without the need for connection establishment procedures

Customer Premises Network (CPN): the customer's private network. In the simplest case, the CPN is just a single end-host or TE

control plane: plane with a layered structure that performs the call control and connection control functions and deals with the signalling necessary to set up, supervise and release calls and connections

data link layer: second layer of the OSI model it provides connectivity between segments of the network (bridging); in addition the data link may perform session control and some configuration

delay variation: delay variation is the difference in delay between successive packet arrivals (of the same flow) at the egress of the network

Differentiated services (Diffserv): services based on statistical (aggregate flows) guarantees and results in "soft" QoS.

NOTE: Using packet markings (code points) and queuing policies it results in some traffic to be better treated or given priority over other (use more bandwidth, experience less loss, etc.)

DS domain: contiguous set of DS nodes which operate with a common service provisioning policy and set of PHB groups implemented on each node

flow: flow of packets is the traffic associated with a given connection or connectionless stream having the same source host, destination host, class of service, and session identification

guaranteed services: using RVSP and integrated services this results in deterministic reservation of network resources and QoS for specific traffic

IP Bearer (service): IP flow between user-network interfaces

management plane: the management plane provides two types of functions, namely layer management and plane management functions:

- **plane management functions:** 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 functions:** performs management functions (e.g. meta-signalling) relating to resources and parameters residing in its protocol entities. Layer Management handles the operation and maintenance (OAM) of information flows specific to the layer concerned

marking: to set the class of service or DSCP of a packet

metering: process of measuring the temporal properties (e.g. rate) of a traffic stream selected by a classifier

NOTE: The instantaneous state of this process may be used to affect the operation of shaper, or dropper, and/or may be used.

multimedia over IP: multimedia calling services, also frequently referred to as MultiMedia over IP (MMoIP), are a generalisation of the VoIP services whereby the communication between end users is enhanced by the use of a variety of media like video, audio, still images and text

network control centre: equipment at OSI Layer 2 that controls the access of terminals to a satellite network, including element management and resource management functionality

Per-Hop Behaviour (PHB): externally observable forwarding treatment applied at a differentiated services-compliant node to a behaviour aggregate

policing: process of discarding packets (by a dropper) within a traffic stream in accordance with the state of a corresponding meter enforcing a traffic profile

Quality of Service (QoS) - IETF definition: the ability to segment traffic or differentiate between traffic types in order for the network to treat certain traffic differently from others. QoS encompasses both the service categorization and the overall performance of the network for each category. It also refers to the capability of a network to provide better service to selected network traffic over various technologies and IP-routed networks that may use any or all of the underlying technologies

Quality of Service (QoS) - ITU definition: QoS is defined as the collective effect of service performances which determines the degree of satisfaction of a user of a service. It is characterized by the combined aspects of performance factors applicable to all services, such as: service operability performance; service accessibility performance; service retainability performance; service integrity performance and other factors specific to each service.

QoS parameters: parameters that will be specified or monitored to ensure QoS

service control function: application layer function that controls the invocation

service levels: end-to-end QoS capabilities of the network which will enable it to deliver a service needed by a specific mix of network traffic

NOTE: The services themselves may differ in their level of QoS.

Service Level Agreement (SLA): agreement between a Service Provider (SP) and its subscriber (or between an SP and an access network operator) characterized by the choice of one data transfer capability and the allocation attribute related to this transfer capability

NOTE: An SLA can also include elements related to traffic policy and availability. It is agreed upon at the initiation of the contract and normally remains the same for all the contract duration.

teleservice: type of telecommunication service that provides the complete capability, including terminal equipment functions, for communication between users according to standardized protocols and transmission capabilities established by agreement between operators

TISPAN: ETSI Project: combines the former ETSI bodies SPAN on fixed network standardization and TIPHON on Voice over IP (VoIP) based networks

Traffic Conditioning Agreement (TCA): agreement specifying classifier rules and any corresponding traffic profiles and metering, marking, discarding and/or shaping rules which are to apply to the traffic streams selected by the classifier

NOTE: A TCA encompasses all of the traffic conditioning rules explicitly specified within a SLA along with all of the rules implicit from the relevant service requirements and/or from a DS domain's service provisioning policy.

transfer capability: capability of transfer of information through a network

NOTE: This term can be used to characterize a telecommunication service or bearer.

user plane: plane with a layered structure that provides user information transfer, along with associated controls (e.g. flow control, recovery from errors, etc)

user: entity that uses the network services requested by the subscriber

3.2 Abbreviations

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

3GPP	Third Generation Partnership Project
ADSL	Asymmetrical Digital Subscriber Line
AF	Assured Forwarding
API	Application Program Interfaces
ATM	Asynchronous Transfer Mode
BB	Bandwidth Broker
BE	Best Effort
BoD	Bandwidth on Demand
BQM	BSM QoS Manager
BRC	BSM Resource Controller
BSM	Broadband Satellite Multimedia
BSM-QRM	BSM QID Resource Manager
COPS-PR	COPS usage for Policy Provisioning
COPS-RSVP	COPS usage for RSVP
CoS	Class of Service
CPE	Customer Premise Equipment
CPN	Customer Premises Network
Diffserv	Differentiated Services (IETF)
EF	Expedited Forwarding
IETF	Internet Engineering Task Force

IntServ	Integrated Services
IP	Internet Protocol
ISP	Internet Service Provider
ITU	International Telecommunication Union
LAN	Local Area Network
LLQ	Low-Latency Queuing
LPDP	Local Policy Decision Point
MAC	Medium Access Control
MIB	Management Information Base
MMoIP	MultiMedia over IP
MPLS	Multiprotocol Label Switching
NCC	Network Control Centre
NGN	Next Generation Networks
NNI	Network to Network Interface
OSI	Open Standards Institute
PDP	Policy Decision Point
PEP	Policy Enforcement Point
PHB	Per-Hop Behaviour
QID	Queue IDentifier
QoS	Quality of Service
RC	Resource Controller
RED	Random Early Detection
RFC	Request For Comments
RSVP	Resource ReserVation Protocol
SCF	Service Control Function
SD	Satellite Dependent
SI	Satellite Independent
SI-SAP	Satellite Independent-Service Access Point
SLA	Service Level Agreement
SP	Service Provider
ST	Satellite Terminal
STQRM	ST QID Resource Manager
STRC	ST Resource Controller
TCA	Traffic Conditioning Agreement
TR	Technical Report
TS	Technical Specification
UNI	User to Network Interface
VoIP	Voice over IP
WRED	Weighted Random Early Detection
DSCP	DiffServ Code Point
NSIS	Next Steps In Signalling (IETF)
SIP	Session Interaction Protocol
SDP	Session Description Protocol
COPS	Common Open Policy Service
GS	Guaranteed Service
CL	Controlled Load
FRED	Flow RED

4 BSM QoS Service Requirements

4.1 General

This clause defines the overall basis for the BSM QoS Service Requirements, which have previously been outlined in TR 102 157 (see Bibliography). Also defined is the way in which these service requirements relate to BSM systems within end-to-end systems. It is assumed that the BSM QoS functional architecture must have the capability of supporting all types of multimedia service (based on IP) and their QoS attributes, although its implementation could be incremental depending on which modules are considered necessary.

The QoS requirements for BSM systems considered here are those which apply to the user accessing IP-based applications and services available in and across heterogeneous networks and to which the BSM system provides access, or plays a role in. These QoS requirements may also be applicable to services in a stand-alone BSM network (for example one in which the BSM acts as a backbone).

The BSM system must offer compatibility therefore with end-to-end network QoS parameters, services and mechanisms which are defined within integrated networks.

The QoS capabilities of IP are considered as the basis of BSM QoS services, without any additional underlying QoS-enabled protocols such as MPLS.

4.2 End-to-End and Bearer QoS

Traditionally, standardization work for public networks (e.g. at the ITU and ETSI) has distinguished between teleservices, which are truly "end-to-end" and operate across terminals and networks (e.g. mouth-to-ear for voice) and network bearer services (e.g. IP flows) that exclude terminals (i.e. operated from UNI to UNI).

In an open, deregulated market it is not always possible to control the user's domestic or corporate installation. So although previously, User QoS specifications have been focused on true end-to-end QoS, it is more appropriate here and in future (or NGN) environments, that QoS for network bearer services is the main focus for network operators.

Nevertheless specifications for the QoS of applications (i.e. teleservices) as seen by the user may often be needed and are handled by application layer functions (e.g. from other service providers) which must convert such QoS specifications into network bearer service QoS specifications.

A bearer service is usually defined as "a type of telecommunication service that provides the capability for the transmission of signals between user-network interfaces", and only involves the OSI layers 1-3 (ITU-T Recommendation I.112, see Bibliography). Therefore for an IP bearer service or flow (i.e. using OSI Layer 3 and below) this definition is applicable between the network edge devices.

As shown in figure 4.1, an end-to-end IP bearer service for the path between UNI interfaces can be defined. The QoS applicable to this IP bearer has been addressed by networking standards such as ITU-T Recommendation Y.1541 (see Bibliography) and TS 123 107 (see Bibliography).

The BSM can be considered as part of the "public" network in the sense that a BSM network operator may want to offer end-to-end services to customers (through SLA's with other network operators), even if a public network does not exist in the traditional network operator sense. The BSM subnetwork QoS contribution would then need to be to considered as part of the network IP bearer.

In an integrated network environment, each network segment would provide its own link layer services (often also called "bearer") such as the BSM Bearer Services (defined at the BSM SI-SAP). The definition of bearer service for link layers therefore applies between NNI's or UNI's and NNI's (rather than only UNI's).

The QoS provided by the bearer service of each network segment or domain must then be taken into account in the specification of end-to-end IP layer QoS, as indicated in figure 4.1. This cumulative effect of network segments applies particularly where end-to-end limits on QoS parameters (such as delay, delay variation, packet loss, error rate) are needed. However, the end-to-end QoS may be less important for lower priority flows, requiring only relative QoS.

These link layer bearer services may be mapped to the IP QoS directly via an interlayer control function or indirectly in the absence of such a control.

BSM Quality of Service is defined by reference to BSM bearer services and associated bearer service attributes (TR 101 984, see Bibliography).

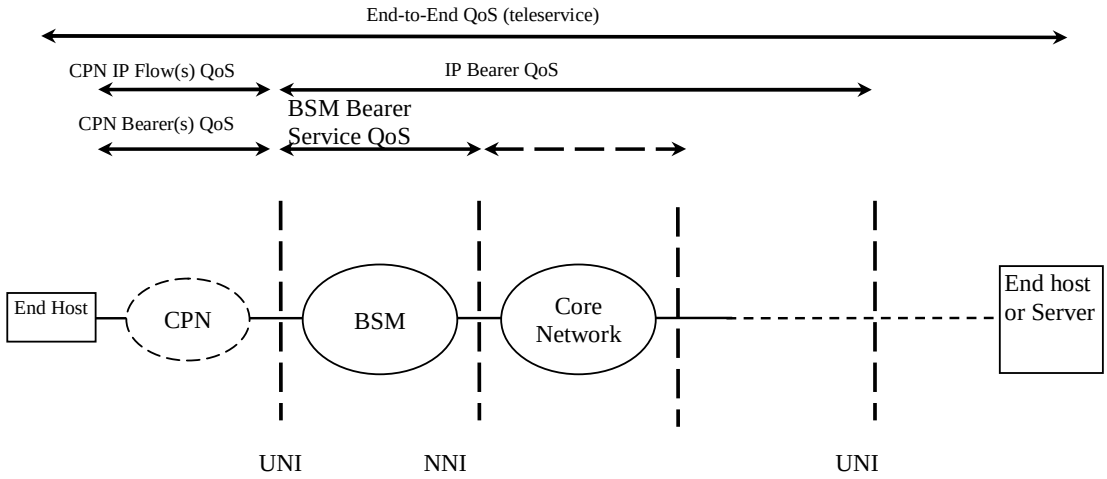


Figure 4.1: End-to-end QoS, and IP QoS and for general IP network scenario

The end-to-end QoS indicated above includes not only the QoS provided by the network, but also that of the CPN or "enterprise". The CPN QoS may be structured into many QoS classes for CPN's of large companies (Cisco: "service provider Quality of Service", see Bibliography).

The IP Bearer shown above, whilst not a constant capacity stream since IP is packet-based and connectionless, nevertheless refers to one flow (amongst possibly a set of flows) of a session set up by a function at higher level or in the control plane.

The BSM system must therefore be capable of satisfying the QoS of IP flows across the integrated network as determined by Service Level Agreements, and this will require mapping the QoS classes of these flows into appropriate BSM bearer classes of service.

Depending on the capabilities of the user equipment and availability of these application services elsewhere in the network, the BSM system may also include application-layer services (proxies) which interact with end-to-end application managers and translate application QoS into IP QoS classes.

4.2.1 Bearer Scenarios for BSM Systems

4.2.1.1 BSM Star Network Topology

In a BSM star network all communication from ST's is with a hub or Gateway station. A BSM Resource Controller (RC) function manages the SI services such as BSM Bearer and BSM IP layer QoS, whilst the Network Control Centre (NCC) manages SD services (OSI layer1 & 2). They could be within a same physical entity but could also be separate. The NCC may or may not be associated with the hub station.

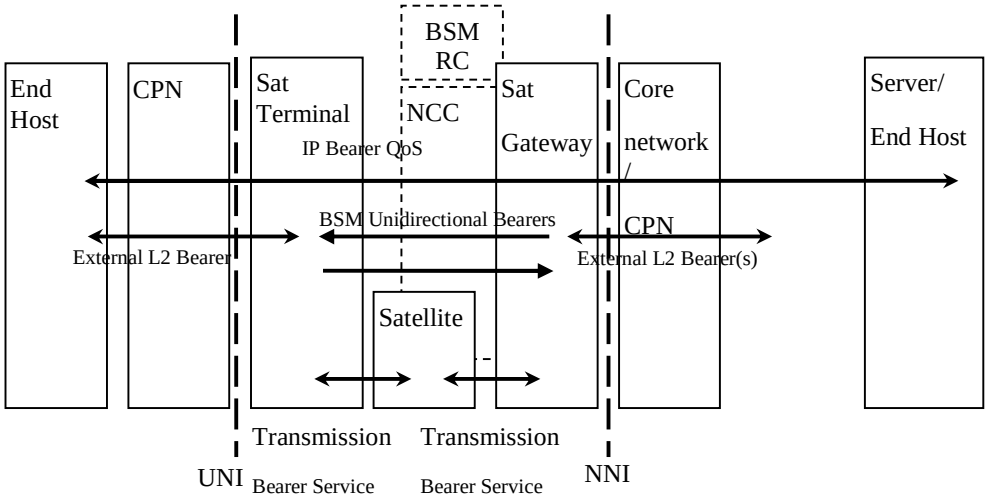


Figure 4.2: BSM Star (Access) Network QoS topology

The BSM unidirectional bearers in the case of the star network are in general different for the forward link and return links.

4.2.1.2 BSM Mesh Topology

In a mesh network there is no conceptual division between inbound and outbound capacity. As for the Star case the BSM RC and NCC are central in setting up calls between sites, but in the mesh case they are not necessarily associated with any particular ST.

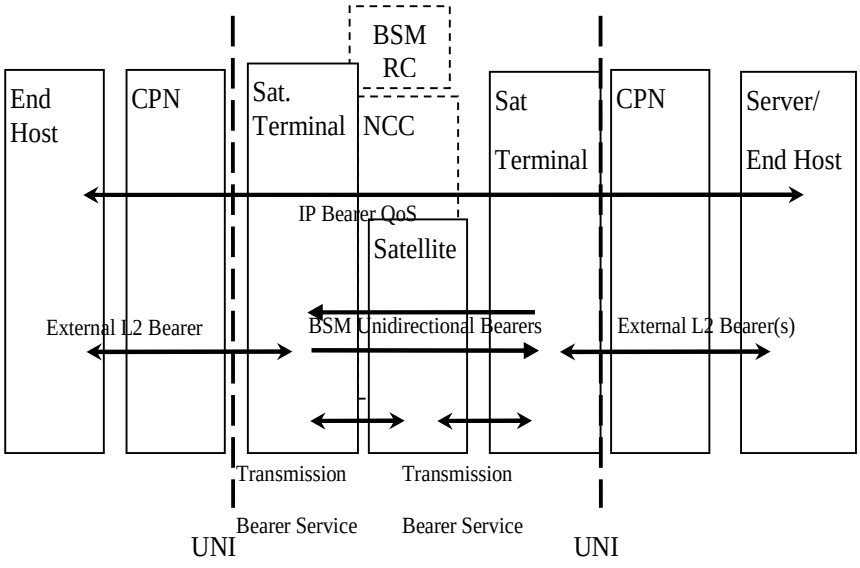


Figure 4.3: BSM Mesh Network Topology

The satellite indicated in the above diagram performs on-board processing (though this type of function is not essential) so that the BSM bearers (indicated as unidirectional - Inbound and Outbound) are regenerated at the satellite. Two bearers are needed (one in each direction) to provide a bi-directional service. In general the QoS of each bearer will be defined separately.

NOTE: A mesh network can also offer a mixed architecture with a star (access) network, where an ST acts as a gateway to a core network. This is not shown in the above for clarity.

4.3 Generic End-to-End QoS Attributes

Leaving aside the best-effort service which is currently prevalent but trivial in terms of QoS, two general models of service assurance can be applied to user and network QoS: guaranteed and relative QoS. The ways in which these are implemented across the network including the BSM subnetwork are fundamentally different and will have an impact on the QoS architecture.

1) Guaranteed QoS refers to a traffic delivery service with numerical bounds on some or all of the QoS parameters. These bounds may be physical limits, or more generally enforced limits such as those imposed through mechanisms like rate policing. The bounds may result from designating a class of network performance objectives for packet transfer for example, and are then assured by performing admission control in the access network via appropriate traffic policing and through put control.

2) Relative QoS refers to a traffic delivery service without absolute bounds on the achieved bandwidth, packet delay or packet loss rates. It describes the circumstances where certain classes of traffic are handled differently from other classes of traffic, and the classes achieve different levels of QoS. This is generally performed by use of separate traffic class queues at the network edge and allocating priorities to the queues.

For example, these models can be related to the IETF IntServ and Diffserv models respectively.

A combination of these models applied to different network segments in an end-to-end path may also be desirable as a compromise to an overall Guaranteed QoS model, due to the scalability difficulties introduced when many end-to-end connections have to be guaranteed. For example a common practical solution is to apply guaranteed QoS to the access network where QoS and sharing of resources is more critical, and to use relative QoS in the core network where capacity is more freely available. Though the overall result will not be absolutely guaranteed and this method can only be applied for suitable traffic classes, the implementation is considerably eased. The implementation of this hybrid solution still needs to be clarified since it needs end-to-end network coordination.

The availability of a Guaranteed QoS model for some services (e.g. VoIP, MMoIP) is considered to be necessary at least over the satellite subnetwork. This is needed for example in the case when the BSM is used as a stand alone end-to-end network, such as a backbone for VPN's, where the end-to-end integrity of QoS can be controlled, such as when corporate LAN's are interconnected over the satellite.

4.4 Multimedia Application QoS Classification

The BSM will need to support potentially a wide range of multimedia applications.

The Multimedia Service Applications to be supported by a system are in general a combination of several service components, each with different QoS requirements. The following service component types can be identified:

- speech: voice telecommunication, focusing on interactive mouth to ear communication;
- audio: telecommunication of audible signals, e.g. music focusing on acoustic fidelity;
- video: telecommunication of motion pictures, focusing on visual fidelity;
- graphics: telecommunication of graphics and still images, focusing on visual fidelity;
- data: telecommunication of data-files, focusing on error-free, and possibly timely, transfer.

Not all Multimedia Service Applications are comprised of all the service components identified above, but rather of one or of a subset.

An overview of a framework for MultiMedia QoS classification is shown in figure 4.4.

QoS classification per service component may take place depending on information like the end-user profile, etc. A service grade makes restrictions like the maximum delay and the maximum packet loss in a certain classification assessment.

Implementation of the transmission facilities for the service component within the BSM system will determine the characteristics and QoS requirements of the resulting traffic flows. Parameters such as choice of codec, packetisation, transport mechanisms (e.g. TCP PEP) deployed will each contribute to the properties of the resulting traffic flow. Therefore the QoS values of the set of service components for a multimedia application need to be chosen and managed accordingly as a set.

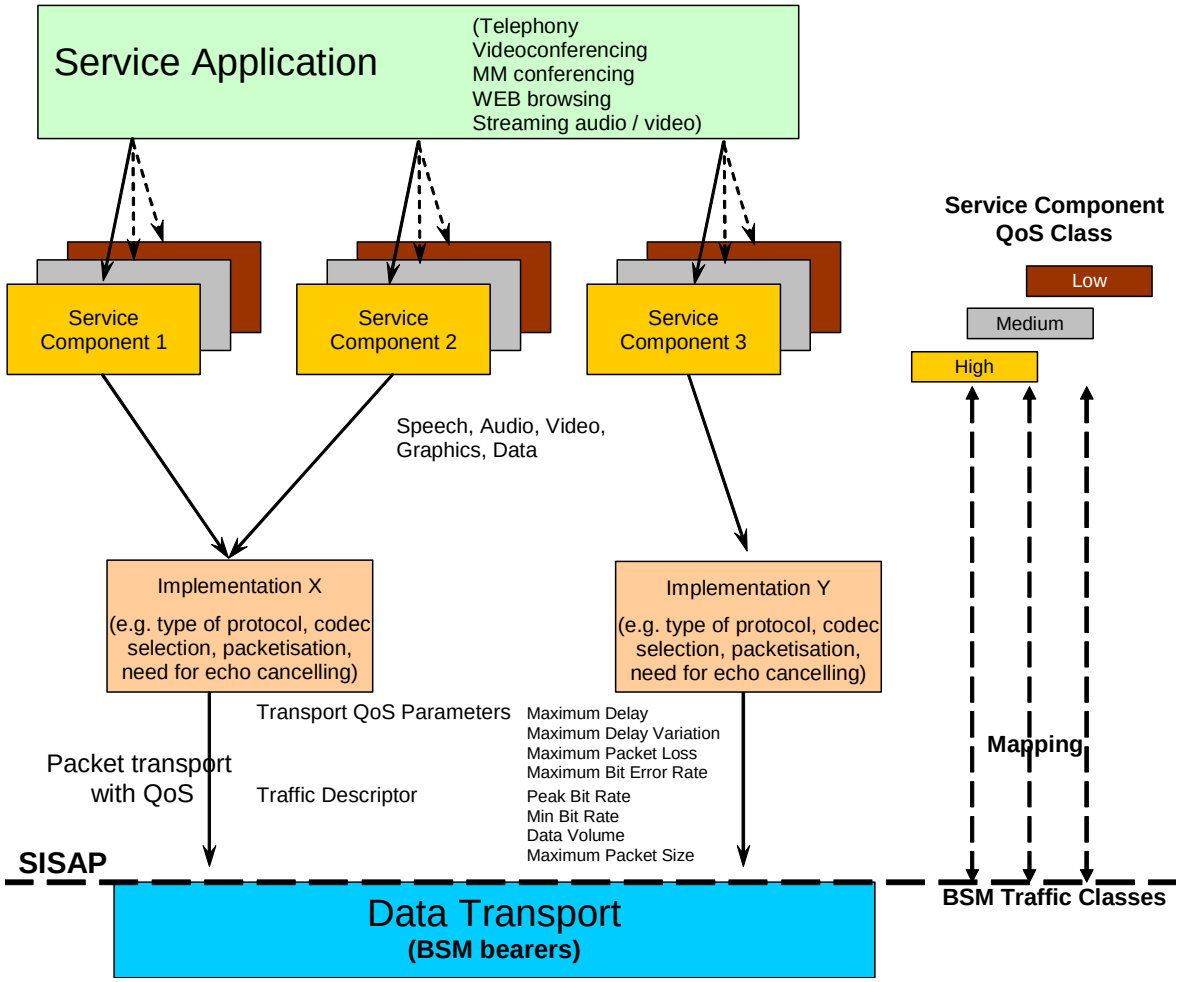


Figure 4.4: Framework for MultiMedia Application Service Set QoS classification

4.5 Relationship between BSM Traffic Classes and IP Traffic Classes

The BSM Traffic Classes are defined at the SI-SAP interface and refer to the CoS of IP packets arriving at this interface. The QoS classification of IP packets entering the BSM system can be defined by a range of parameters (such as Diffserv code points, IP type field, etc.) and therefore the class attributes of IP packets need to be mapped to and from the BSM traffic classes. This is an area of interest for future work.

Among the possibilities, for example, is that different DS domains may use a different meaning for the same DSCP, and the mapping of 64 potential DSCP's into a lesser number of classes may not be a trivial task if the characteristic of each class is not known.

This mapping is ideally done by functions placed at the BSM network edges, i.e. in the ST's. The mapping between classes is not a trivial function and needs to be carefully defined.

An example of end-to-end Classes is given in annex A.

BSM Traffic Classes are shown in annex B.

QoS traffic handling at the SI-SAP shall be as defined in [3] in which a set of queuing identifiers (QID's) is also defined; a QID enables the data (IP packets) to be allocated to a queue, and then to be policed and transported properly across the BSM system.

4.6 Service Requirements - Conclusions

The BSM QoS Architecture should support the following service requirements:

1) compatibility with end-to-end IP network QoS parameters, services and mechanisms within integrated networks;

2) satisfying the QoS requirements of IP flows across the integrated network as determined by Service Level Agreements;

3) support control of both relative QoS and guaranteed QoS;

4) the use of BSM traffic classes to define the QoS properties for the transport of IP packets across the BSM subnetwork;

5) the mapping of the QoS attributes of IP packets to and from SI-SAP QoS attributes at the BSM subnetwork edges.

5 BSM QoS Functional Architecture Requirements

5.1 General

The aim of this clause is to describe the "global picture" of where the BSM system fits within end-to-end QoS-enabled networks, including the functions and interfaces required.

5.2 QoS Architecture Approach for IP Networks

Overall approaches to QoS architectures in emerging and future IP-based networks have been described in RFC 2990, (see Bibliography), 3GPP (see Bibliography), ETSI TISPAN (NGN) (see Bibliography), ITU-T (see ITU-T Recommendation Y.1291 in Bibliography), ADSL Forum (see Technical Report 059 in Bibliography), etc.

One of the main common characteristics of these approaches, is the uncoupling of services and networks, allowing services and networks to be offered separately and to evolve independently. Therefore in the architectures described there is a clear separation between the functions for services and for transport, and an open interface is provided between them. Provisioning of existing and new services can then be independent of the network and the access technology.

In emerging networks (such as NGN's) there is increased emphasis by service providers on service customisation by their customers by means of service related APIs (Application Program Interfaces) in order to support the creation, provisioning and management of services. In such networks the functional entities controlling policy, sessions, media, resources, service delivery, security, etc. may be distributed over the infrastructure, including both existing and new networks. When they are physically distributed they should communicate over open interfaces.

The basic approach is to differentiate between the Application (Service) and Transport Strata as follows.

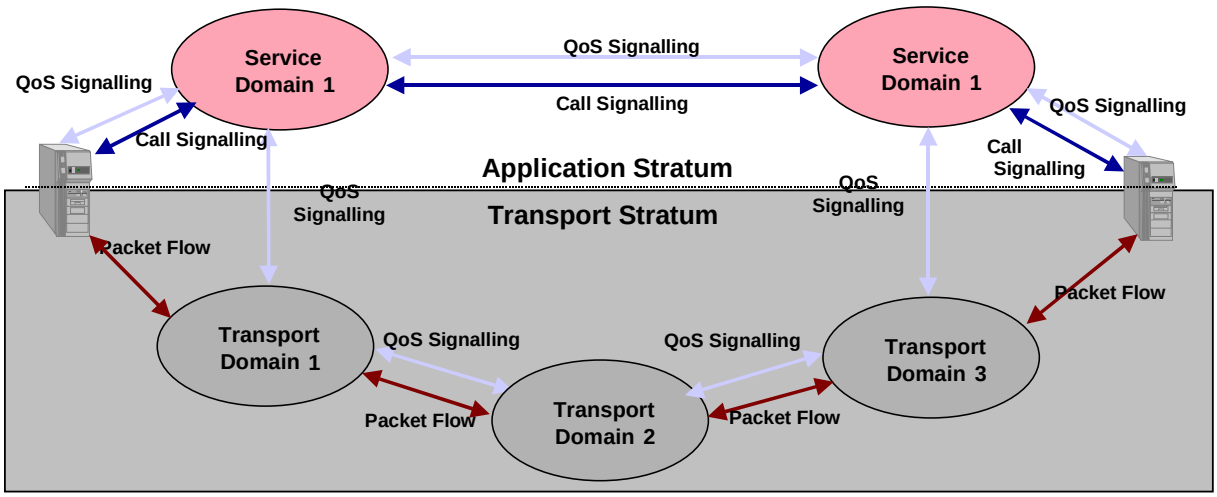


Figure 5.1: NGN Application & Transport Strata

The Application Stratum provides the service to users. Service is requested by user/call signalling protocols (e.g. H225, +H245, SIP +SDP, H248).

The signalling allows description of :

- the user end-points of the session;
- QoS parameters (codec, frames per packet, frame size, jitter buffer delay, FEC, mean delay variation, packet loss);
- other service related parameters.

The Transport Stratum provides a packet-oriented transport service and the desired network QoS. The QoS is requested by QoS signalling protocols (e.g. RSVP, COPS, NSIS). QoS signalling can be exchanged with user endpoints and/or the application plane.

The main assumptions are the following:

- the media path may cross several transport domains;
- transport domains may support different QoS mechanisms and policies;
- routing of calls between transport domains will be under the control of the application plane (one or more SPs);
- routing of calls within the transport domains will be independent of the application stratum.

5.3 QoS Network Building Blocks

To offer QoS services outlined above in a complete and efficient way can be complex, and can involve multiple inter-related aspects. For example, in case of network resource contention or congestion, to maintain the expected service response requires a variety of functions working at different time-scales, ranging from careful network planning based on traffic patterns over a long period (grouped in the Management Plane) to differential resource allocation and admission control based on the current network load conditions (in the Control Plane).

The range of mechanisms involved in QoS can be considered as a set of **building blocks**, or functions which can be combined in different ways to provide different overall objectives (e.g. type of network, or for guaranteed or relative QoS). These building blocks may be classified in the Management, Control and Data Planes as follows.

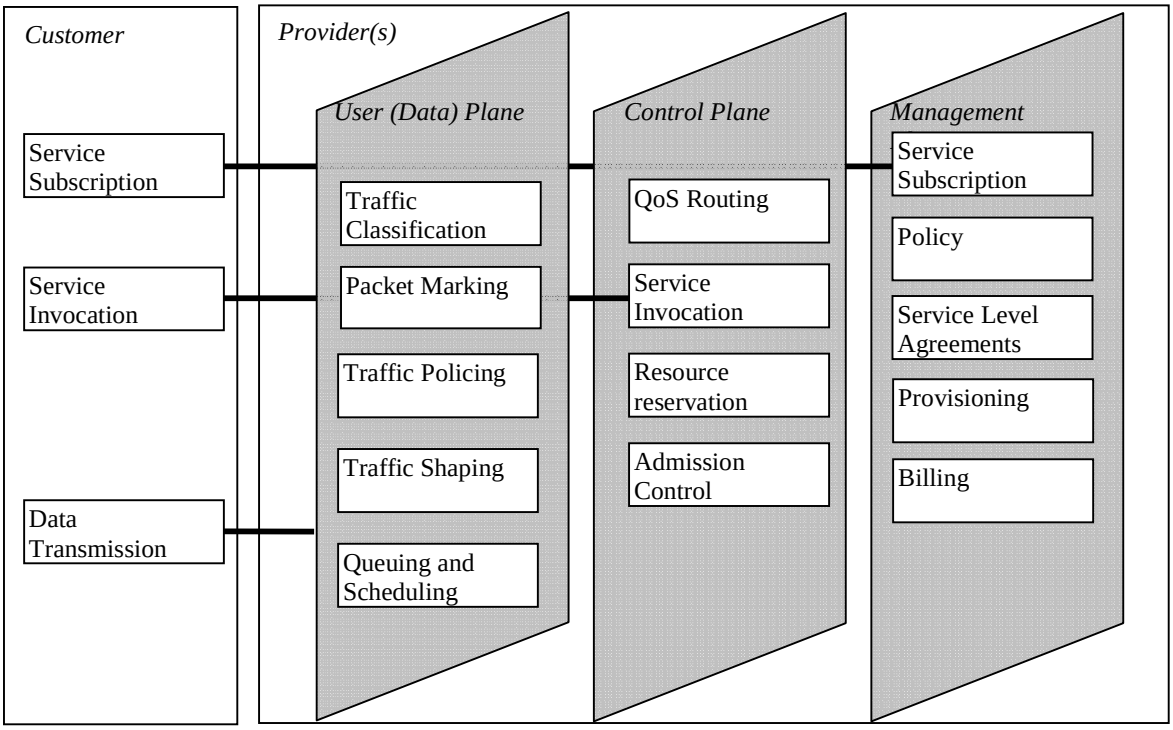


Figure 5.2: Functional/Architectural Framework for QoS provision

The above figure is based on that in ITU-T Recommendation Y.1291 (see Bibliography).

The above framework is intended to be application independent.

A comprehensive QoS solution typically employs multiple building blocks across the Management Plane, Control Plane and Data Plane, but practical implementations may require only a subset of the functions.

The functions in each of these planes are described in annex C.

5.4 Interactions between building blocks

QoS parameters need to be exchanged between the various building blocks. These parameters include transaction performance at the packet level (e.g. delay and packet loss) and service reliability/availability expectations in the form of traffic priority levels for specific network functions such as admission control and traffic restoration. Examples for mechanisms to convey these parameter values are signalling and database lookups and include:

- QoS signalling: signalling of QoS parameter requirements per service/flow between functional blocks;
- Call Signalling: service invocation, resource reservation;
- Policy control: parameters for admission control, policing, marking, etc.

5.4.1 Options for QoS Architectures

One of the main options for QoS architectures is whether QoS is a per-application service or a network (Transport Stratum) signalled service (or both):

1) Application-based QoS leads to extension of the QoS architecture into some form of Application Program Interface (API), so that applications can negotiate a QoS response from the network and alter their behaviour accordingly. Examples of this approach include GQOS, and RAPI (see The Open Group: "Resource Reservation Setup Protocol API, C809" in Bibliography), and the ESA TRANSAT project (see ICN 2005: 4th International Conference on Networking (April 17-21, 2005) Proceedings: "Quality of Service Solutions in Satellite Communication" in Bibliography).

2) For network-signalled QoS, any application could instead have its traffic carried by some form of QoS-enabled network services by changing the host configuration (e.g. packet marking), or by changing the configuration at some other network control point, without making any explicit changes to the application itself.

The strength of the latter Transport Stratum approach is that there is no requirement to alter substantially the application behaviour, as the application is itself unaware of the administratively assigned QoS. The weakness of this approach is that the application is unable to communicate potentially useful information on its nature to the network or to the policy systems that are managing the network's service responses. In the absence of such information the network may provide a service response that is far superior or inferior to the application's requirements. An additional weakness is for the type of applications that can adapt their traffic profile to meet the available resources within the network, since any network availability information is not passed back to the application.

Application Service invocation is therefore assumed to be controlled by either:

a) a function in the service provider's network (called here the Service Control Function (SCF)) which manages the end-to-end service, including any QoS requirements; or

b) specific session and QoS signalling (e.g. SIP/SDP).

In the first case the user's CPE initiates the service by sending a request to the SCF, and in the second by issuing QoS requests to the network. These cases are further illustrated below.

5.4.2 QoS Signalling

For guaranteed services, QoS for multimedia applications can be requested from the network in several ways. The CPE, SCF's, BSM Managers and BSM ST's (edge routers) need to cooperate to provide QoS in the access network based on resource and admission control.

For relative QoS, the QoS parameters available to a customer are normally determined by an SLA, and any variation required should be handled by SLA renegotiation (or a bandwidth broker). If the network is incapable of meeting the service request (e.g. if the network itself changes state and is unable to meet the cumulative traffic loads admitted by the ingress traffic conditioners), neither the ingress traffic conditioners, nor the applications, are informed of this failure to maintain the associated service quality and the request simply will not be honoured. There is no requirement for the network to inform the application of this situation, and the application must determine if the service has not been delivered.

In future, network availability as well as resource requirements at least for high reliability services in DiffServ should be possible to be signalled by means of NSIS (see RFC 2990 in Bibliography) or RSVP.

5.4.3 QoS Call Signalling and Policy Control Scenarios

The following scenarios illustrate the interactions between QoS and call signalling and policy control.

Different QoS negotiation scenarios identified are driven by the capabilities of CPE's and service providers corresponding to different business models and to different QoS network technologies, summarised as follows:

the CPE's capability to request QoS:

1) not at all (implies default case of relative QoS availability only; the CPE may or may not be able to mark packets);

   a) via an application protocol, or an API (see clause 5.4.1);

   b) via dedicated IP QoS signalling for requesting resources;

2) the service provider's ability to offer services either alone or associated with different supporting network QoS policy mechanisms.

Depending on these capabilities the following main QoS negotiation scenarios have been identified, (see IST MUSE Project in Bibliography):

1) service provider-oriented model (Proxied QoS with policy-push):
The service provider's SCF from which a specific service is requested is responsible for appropriate QoS and policy requested from the network providers before accepting the request. The CPE has no involvement in explicit QoS signalling, which is contained solely within the network providers domain.

2) User-oriented model (User-requested QoS with policy-pull):
The CPE can initiate explicit IP QoS requests without service authorisation from the SCF in advance. The authorisation for the IP QoS request is performed directly by the network on receipt of that request.

3) Application-signalling-based models:

   a) (User-requested QoS with policy-push-pull) The CPE can initiate explicit IP QoS requests through network QoS signalling (e.g. RSVP, NSIS), but QoS service authorisation is still needed from the SCF in advance;

   b) The CPE can initiate IP QoS requests through application layer signalling (e.g. SDP/SIP).

These scenarios are described in more detail below.

5.4.3.1 Proxied QoS with policy-push

In this case the client's terminal does not itself support native QoS signalling mechanisms. It requests an application-specific service by sending a "Service Request" (e.g. SIP Invite) to the SCF. This request contains no QoS service parameters. It is then the SCF's responsibility to determine the QoS needs ("proxied" QoS) of the requested service and to request network authorization from the Resource Controller(s) which then requests resource reservation to the Access network (and to Core network, etc.). When the service is authorised, policy (admission control, packet marking, traffic control) is sent from the Resource Controller. Depending on the relative QoS or guaranteed QoS model, the DiffServ marking and enforcement of admission control decisions, via throughput control and traffic conditioning, are properly set in the IP edge (i.e. policy decisions are pushed to the Policy Enforcement Point (PEP) at IP edge or ST).

Policy can also be pushed to the gateway ST. Policy push is usually performed via the BSM-RC and NCC (when it involves layer-2 reconfiguration), as shown in figure 5.3.

This model is typical of xDSL (see Technial Report 059 in Bibliography).

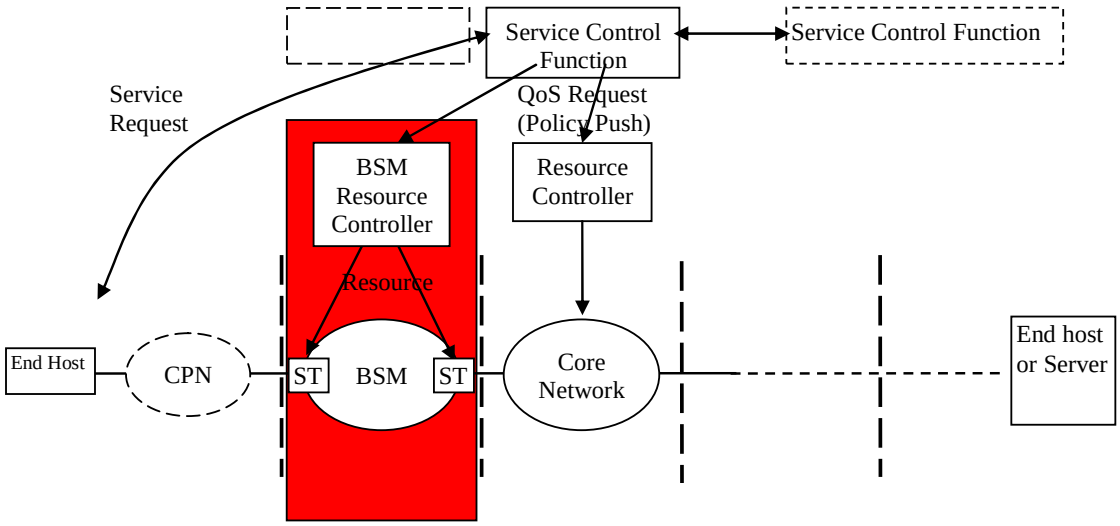


Figure 5.3: Proxied QoS with policy-push

This scenario has the advantages of not requiring resource reservation signalling capabilities in the user terminal and no special protocol for the service session requests. The drawback of this approach is the necessity to always go through the Service controller for any service request, including bandwidth reservation changes during a session.

This scenario supports single-phase resource reservation or two-phase resource reservation:

1) the network enables immediate activation and usage of network resources by the end-user without being dependent on resource availability;

2) the Service controller asks for network QoS resources to be authorised and reserved. Once these resources have been reserved, the Service controller continues its dialogue with the user concerning the service. This two-phase reserve/commit model guarantees that access-network resources are available before offering service to the user.

5.4.3.2 User-requested QoS with policy-pull

In this case there is little or no signalling relationship between the service provider and network provider. The user's terminal is capable of sending QoS Request over Layer 3 QoS signalling for its own QoS needs (policy pull). Authorisation for the IP QoS request is obtained "on the fly" at the time the QoS request is actually signalled and does not require prior authorisation. No communications by the user with the Service Controller is needed prior to making the QoS request to obtain the corresponding authorization. This avoids the need for the Network Resource Controller to maintain awareness of the relationship between end-users and their corresponding PEPs.

The CPE is able to establish QoS reservation end-to-end since the IP QoS signalling proceeds hop-by-hop. It requires the CPE to support Layer 3 QoS signalling.

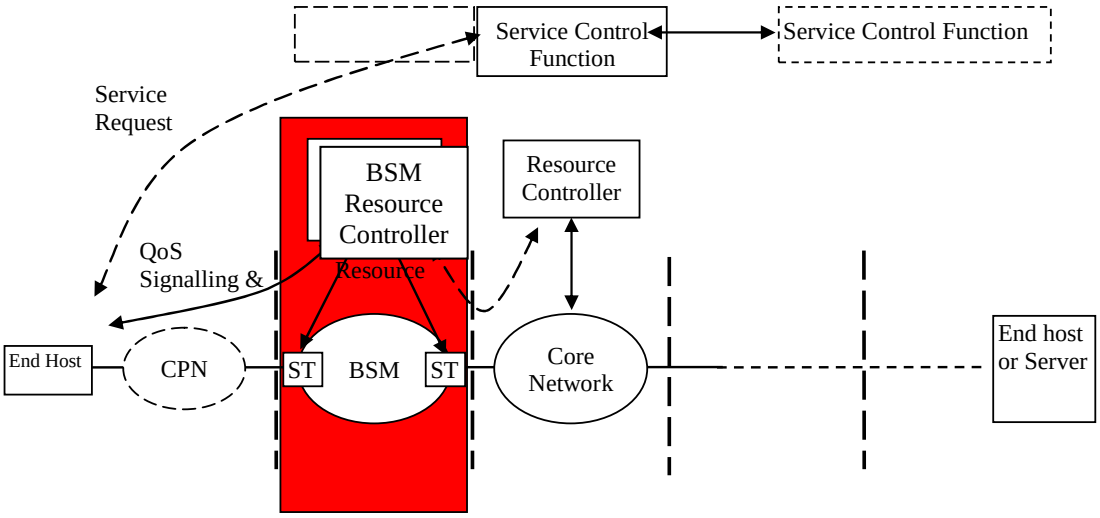


Figure 5.4: User-requested QoS with policy-pull

5.4.3.3 User-requested QoS with policy-push-pull

The difference with the previous case in clause 5.4.3.2 is that the user terminal is capable of signalling and managing IP QoS resource requests but requires prior authorization to do so from the Service controller. The user requests an application-specific service by sending a "Service Request" to the Service controller. The Service controller is in charge of determining the QoS needs of the requested service and of requesting the network authorization from the Network resource controller (policy push).

The terminal then uses network signalling to request resource reservation and commitment (policy pull). This request could be managed in the Access Network with authorisation of the Network Resource Controller.

This model has the ability to reserve end-to-end QoS since the IP QoS signalling proceeds hop by hop.

(This model is typical of UMTS networks).

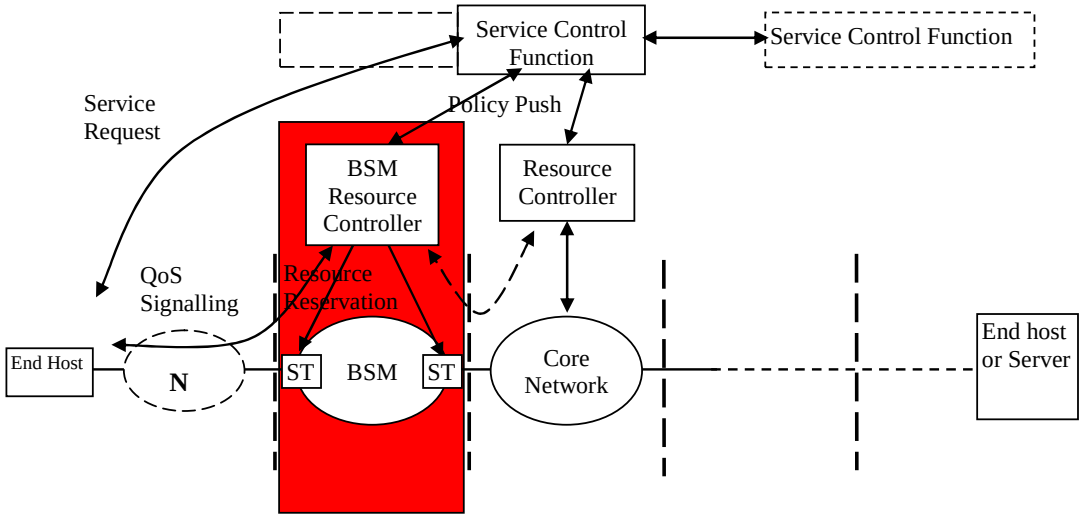


Figure 5.5: User-requested QoS with policy-push-pull

5.4.3.4 User-originated application layer QoS signalling

The user requests a service from a service provider via application layer signalling, and the QoS requirements are included in the signalling. The QoS requirements can be reserved end-to-end since the application layer signalling proceeds hop by hop through service domains. The SCF extracts the QoS requirements interprets them and passes the QoS request to the IP layer. The Service and Network Providers retain control of QoS authorisation in this model.

5.5 Policy Control

As indicated in clause 5.4 the main options for policy control of network elements (the ST edge router and any other local routers) are:

1) locally provisioned policy;

2) outsourced policy to a policy server.

These options correspond to the policy push and pull models respectively. The COPS protocol is commonly used for transfer of policy.

There are complex design tradeoffs in the use of the push model (server policy push) or the pull model (client/local router policy request). Some of these are summarised as follows:

For the Push Approach:

- local policy reduces the delay encountered with the COPS REQUEST/ DECISION exchanges during outsourcing policy control to the policy server;
- it requires new behaviour in the network elements to accommodate configuration of local decision policy;
- requires network topology knowledge to determine the network elements needing the local decision policy.

The push model offers the lowest possible call setup delay, but requires protocols and implementation in the network.

COPS-PR (see RFC 3084 in Bibliography) is the protocol that is used when policy decisions are "pushed" from the PDP to PEPs. In this provisioning model PDP can send policy decisions to PEPs without having specific request from PEP.

For the Pull Approach:

- additional messaging resulting from the edge router outsourcing policy control to the policy server;
- does not require any additional protocol work, uses existing COPS protocols;
- does not require knowledge of network topology.

COPS-RSVP is the protocol that is used when policy decision is "pulled" from the PDP. When an RSVP message requiring a policy decision is received by PEP the relevant RSVP objects from the message are put into a COPS Request message, which is sent to PDP. The PDP determines what to do with RSVP message and sends a COPS Decision message back to the PEP, in response to the Request (see RFC 2749 in Bibliography).

The pull approach is therefore preferable for faster deployment and lower development costs for QoS in the network.

5.6 BSM Global QOS Architecture

The BSM Global QoS functional architecture, including the relationship of the BSM with QoS protocol layering and with the rest of the network, is illustrated in the following diagram. This diagram is based on the Functional/Architectural Framework of figure 4.4.

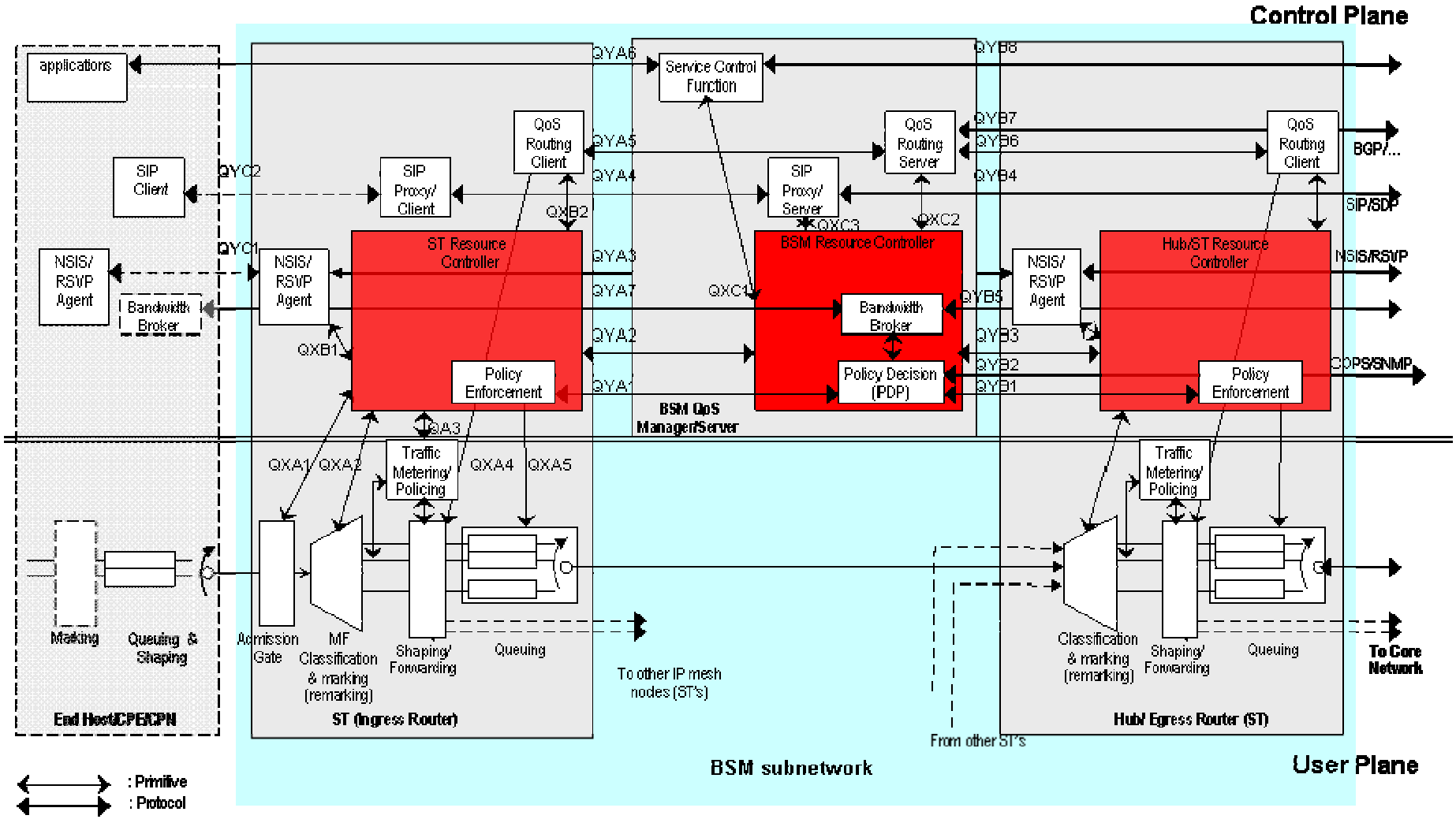


Figure 5.6: BSM QoS Functions in the IP layer and higher layers (one data direction shown)

Figure 5.6 shows the range of possible functions involved in QoS and their functional partition between Control and User Planes (Management functions are not shown for clarity since they are more implementation-specific).

Two main kinds of message flows between functional blocks are indicated in the diagram: primitives between protocol layers, and secondly peer-to-peer protocols. Note that the peer-to-peer protocols are shown as horizontal lines for clarity, though in reality they are transported via the user plane.

The Interfaces are listed below.

Interface Label	Description	Example Protocol
QXA1	Admission policy	primitive
QXA2	Classification policy	primitive
QXA3	Traffic conditioning policy/traffic measurement	primitive
QXA4	QoS Routing/Forwarding Table	primitive
QXA5	Queuing and scheduling policy	primitive
QXB1	IntServ & DiffServ reservation	primitive
QXB2	QoS routing Table	primitive
QXC1	Service Resource Request/Response	primitive
QXC2	QoS routing Table	primitive
QXC3	SIP service Request/Response	primitive
QYA1	Policy Request/Response Internal	COPS/SNMP
QYA2	Resource Control	BSM defined
QYA3	NSIS/RSVP resource reservation signalling	NSIS/RSVP
QYA4	SIP/SDP signalling	SIP/SDP
QYA5	QoS Routing Internal	IGP/BGP
QYA6	Service Request/Response	BSM defined
QYA7	SLA negotiation	TBD
QYB1	Policy Request/Response Internal	COPS/SNMP
QYB2	Policy Request/Response External	COPS/SNMP
QYB3	Resource Control	BSM defined
QYB4	SIP/SDP signalling	SIP/SDP
QYB5	SLA negotiation	TBD
QYB6	QoS Routing Internal	IGP/BGP
QYB7	QoS Routing External	IGP/BGP
QYB8	Service Request/Response	BSM defined
QYC1	NSIS/RSVP resource reservation signalling	NSIS/RSVP
QYC2	SIP/SDP signalling	SIP/SDP

5.6.1 ST

As an edge router the ST should contain some of the key functions for QoS services, since it needs to provide the first point of support to clients requesting service. An ST can be of two principal types:

1) a user-access ST (for star or mesh networks);

2) a Hub ST for star networks.

Both types may need to act as ingress and egress edge routers towards external networks, depending on where the boundary of the service provider's network is drawn. Therefore the overall range of functions needed for each ST type may be the same, but their traffic scales and their number of network interconnections will be different, since the Hub ST acts as a concentrator for traffic in the BSM return links.

The range of QoS-related functions is indicated in figure 5.6.

For operation for guaranteed QoS under the IntServ model, the end-hosts signal their QoS needs to the global network via the ST which, after resource set-up, performs per-flow handling and admission control.

For relative QoS operation under the DiffServ model, provisioned QoS is generally adopted by setting up network elements with sufficient resources to service multiple classes of traffic, with varying QoS requirements. However dynamic resource reservation can also be performed by means of a bandwidth broker or RSVP signalling.

The range of ST QoS functions for the User, Control and Management Planes are described in annex C.

5.6.1.1 Diffserv operation

The BSM shall be compatible with the IETF Diffserv model as defined in [3]. For the DiffServ model the ST's can typically be categorized as Boundary nodes, as they may act as demarcation points between the DS-Domain and a non-DS-aware (e.g LAN) network. The Hub ST may, in addition, act as a DS boundary node connecting two DS Domains together, if the BSM does not form part of a larger DS Domain (figure 5.6).

Typically, the Boundary node performs traffic conditioning, unless the traffic is pre-marked and conditioned by the traffic sources directly, or by nodes in the upstream network (e.g. CPN). The SLA between an upstream network and a downstream DS domain may specify packet classification and re-marking rules and may also specify traffic profiles and actions to traffic streams which are in- or out-of-profile. The Traffic Conditioning Agreement (TCA) between the domains is derived from this SLA.

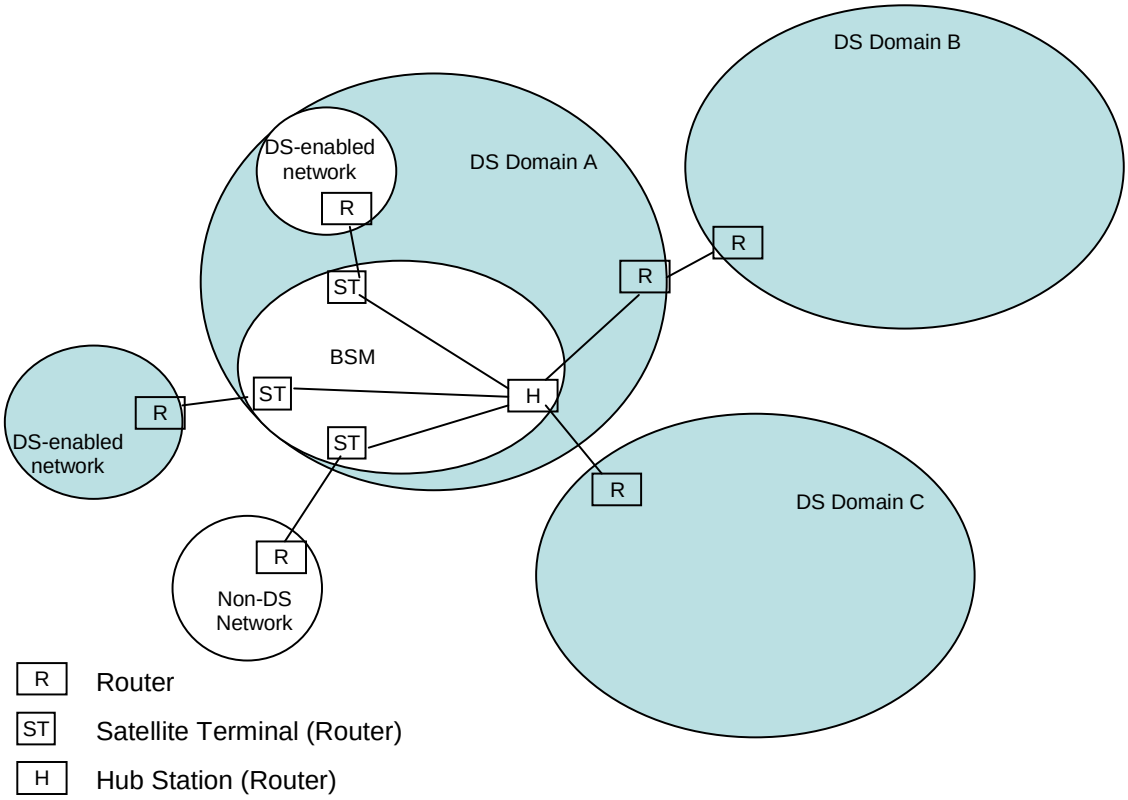


Figure 5.7: Scenarios for ST Boundary nodes (Star Network) within DiffServ Domains

Traffic streams may be conditioned on either side of a boundary (the DS egress node of the upstream domain or the DS ingress node of the downstream domain). However, a DS ingress node should assume that the incoming traffic may not conform to the TCA and must be prepared to enforce the TCA in accordance with local policy.

In traffic conditioning the incoming packets may be classified into pre-defined aggregates, metered to determine compliance to traffic parameters (and determines if the packet is in profile, or out of profile), marked appropriately by writing/re-writing the DSCP, and finally shaped (buffered to achieve a target flow rate) or dropped in case of congestion. By this means as a packet leaves the ST Ingress router and enters the BSM network, PHBs will be enforced, depending on the packet marking with the appropriate DSCP. However, the extent of traffic conditioning required is dependent on the specifics of the service offering, and may range from simple DSCP re-marking to complex policing and shaping operations. For example in a small sub-network only policing of traffic may be necessary, with no requirement for classification.

If the ST is employed as a DS Internal node in a DS Domain then it must still enforce the appropriate PHB for the Domain by employing policing or shaping techniques, and sometimes re-marking out-of-profile packets, depending on the policy or the SLA.

5.6.2 Resource Management

The approach adopted to the BSM QoS architecture is based on centralised control and management of the BSM subnetwork through a Server entity called the BSM QoS Manager (BQM) (see TR 102 157 in Bibliography). (Like typical servers the BQM can consist of several physical entities). The ST's, as network edge devices, are responsible for traffic processing and policy enforcement at the ingress and egress, but they should be controlled from the BQM. The BQM should contain all the necessary functions to manage QoS for all layers above the SISAP in both Management and Control Planes. The BQM interacts with equivalent local functions in the ST's.

The control and management functions below the SISAP (for connection control, bearer set up, BSM QoS etc.) are usually also centralised in the NCC, which may be closely associated with the BQM.

Many of the functions in the BQM are standardised functions such as those for signalling (RSVP/NSIS or SIP Proxy/SDP), but others specific to the BSM, such as those for managing the BSM's global IP and SIAF layer resources, are allocated to a functional entity called the BSM Resource Controller (BRC).

If the BSM is an independent administrative domain then the BQM acts as an self-standing PDP (Policy Decision Point). If the BSM is part of a wider domain then the BQM may act as a Local PDP (LPDP) to which the policy database is downloaded from the primary PDP, in order to avoid scalability problems and excess management traffic in making many policy requests to a PDP located outside of the BSM. The BQM PDP makes policy decisions using service-based policy rules, and communicates these decisions to the Policy Enforcement Points (PEPs) in the Resource Controllers (client) of the ST's (STRCs).

NOTE: The discrepancy of actual versus the predicted resource availability which occurs in practice (due to traffic variation) is a major issue. Care should therefore be given in the resource management architecture to use the most current resource information to make link, system and other resources available for the requesting application.

5.6.2.1 BSM Resource Controller

The BRC is intended to act as a server function for the local Resource Controllers in the ST's (STRCs) which act as clients.

The BSM Resource Controller should:

- have a view on available network resources;
- have a view on the allocation of IP addresses;
- have a view on the use of network resources;
- control admission of new IP flows based on resource availability;
- control policing functions of ST's to allow traffic to use reserved resources.

The BRC may perform several functions such as policy control and admission control. The BRC should choose the local policy to be applied to a request from the SCF (or from other triggers) based on, for example, the application type, requested priority level, and the indicated resource reservation service class.

The BRC contains the Bandwidth Broker (BB) which (re-)negotiates SLAs to modify resources between domains.

The BRC may also contain a PDP (if not included within the BB) which sets policy in PEPs or makes policy decisions on request from PEPs in the ST Resource Controllers and returns the results in the form of policy.

An example of a simple admission control scheme for flow-level dynamic QoS provisioning under the centralised BRC model is as follows. When a new flow arrives at an ST edge router requesting a certain amount of bandwidth to satisfy its QoS requirement, the flow reservation set-up request is forwarded by the ST to the BRC. The BRC then applies an admissibility test to determine whether the new flow can be admitted. the BRC examines the path QoS state and determines whether there is sufficient bandwidth available along the path to accommodate the new flow. If the flow can be admitted, the BRC updates the path QoS state database and link QoS state database (as well as the flow information database) to reflect the new bandwidth reservation along the path. If the admissibility test fails, the new flow reservation set-up request will be rejected, and no QoS information databases will be updated. In either case, the BRC will signal the ingress edge router of its decision. For a flow reservation tear-down request, the BRC will simply update the corresponding link state database and path state database (as well as the flow information database) to reflect the departure of the flow.

As indicated above, the BRC should also maintain a number of management information bases (MIB) for the purpose of QoS control and management of the BSM domain, for example:

- topology information base (information that the BRC uses for route selection and other management and operation purposes);
- policy information base (contains policies and other administrative regulations of the domain);
- link QoS state information base.

Other MIBs that may be used by the admission control function are:

1) Flow Information Base.

This MIB contains information regarding individual flows such as flow ID, traffic profile, service profile (e.g. end-to-end delay requirement), route ID and QoS reservation associated with each flow. Other administrative (e.g. policy, billing) information for a flow may also be maintained here.

2) BSM network QoS State Information Bases.

These MIBs maintain the QoS states of the BSM domain, and are the key to the QoS control and management of the domain. The network QoS state information can be represented in two-levels using two separate MIBs: *path QoS state information base and ST node QoS state information base*. These two MIBs are as follows:

- **Path QoS state information base** maintains a set of paths (each with a route ID) between the STs (and other routers) of the domain. Associated with each path are static parameters and dynamic QoS state information. Examples of static parameters are:
  - number of rate-based schedulers and delay-based schedulers;
  - sum of the propagation delay along the path;
  - maximum permissible packet size.

  The dynamic QoS parameters are:
  - the set of flows traversing the path (in the case of per-flow guaranteed delay services);
  - or the set of delay service classes and their aggregate traffic profiles (in the case of class-based guaranteed delay services);
  - a number of QoS state parameters regarding the (current) QoS reservation status of the path (such as the minimal remaining bandwidth along the path, delay parameters currently supported along the path).
- **ST QoS state information base** maintains information regarding the ST routers in the domain. Associated with each router is a set of static and a set of dynamic parameters representing its QoS state. Examples of static parameters are:
  - the scheduler type(s) (i.e., rate- or delay-based), its error term(s);
  - propagation delays to its next-hop routers;

- configured total bandwidth and buffer size for supporting guaranteed delay services;
- a set of delay classes and their associated delay parameters;
- and/or a set of pre-provisioned bandwidth and buffer size pairs for each delay class supported by the scheduler.

The dynamic router QoS state parameters include:

- the current residual bandwidth at the scheduler;
- a list of delay parameters associated with flows traversing the scheduler.

5.6.3 Guaranteed and Relative QoS Coexistence

IntServ and DiffServ can co-exist as two models for End-to-End QoS. The DiffServ Domain should pass the IntServ reservation requests transparently, while providing its own policy-based PHBs through it. Devices outside of the DS-Domain reserve bandwidth using RSVP.

Both IntServ and DiffServ models can be driven from a policy database, using COPS for example. The ST then acts as a Policy Enforcement Point (PEP) under control of the policy server (PDP).

5.6.4 QoS Routing

Neither IntServ nor DiffServ models combine signalling/provisioning with the routing process. There may exist a path in the network that has the required resources, even when RSVP/DiffServ fails to find the resources. True QoS, with maximum network utilization needs the marriage of traditional QoS and routing. A solution could be to use traffic engineering and MPLS.

5.6.5 SIP Proxy

The SIP proxy server acts on the behalf of and provides services to all SIP clients in the access network or the administrative domain. Clients requesting call setup have to be first registered with the SIP server, before obtaining authorization for QoS supported calls. After registration with the SIP server, the server may handle all call requests to/from that client. This does not exclude however the client performing direct client-client call setup without the benefits of any SIP server. Such direct client-client call setups can be faster and may be desirable for special services. Clients that are not registered and authorized for direct calling, cannot have the QoS benefit via the support from the SIP and policy servers.

5.7 QoS Architecture Requirements - Conclusion

The BSM QoS Architecture should support the following functional requirements:

1) Compatibility with NGN principles of uncoupling of services and networks.
2) Capability to include as many of the QoS functional building blocks as needed to achieve the required overall network performance i.e. modular concept.
3) Capability to implement the range of methods for QoS signalling (request, initiation and negotiation) including:
   - service provider-oriented model (Proxied QoS with policy-push);
   - User-oriented model (User-requested QoS with policy-pull);
   - Application-signalling-based models.

4) A client/server resource management architecture, where the client side resides in the STs and the server side provides centralised functions for control and management for the complete BSM network as defined in the generic BSM architecture (see TR 102 157 in Bibliography).

NOTE: The server side is modelled as a centralised function but this function may be implemented as one location or spread over multiple locations (i.e. this centralised function may be physically distributed).

6 BSM QoS Functional Architecture Definition

6.1 Scope

No standardized or common approach to network architecture for end-to-end QoS provision to applications exists at present.

Various approaches to QoS provision can be proposed based on varying complexity and performance. The objective of this clause is to identify several BSM QoS Scenarios and associated architectures. These architectures are required to be compatible with the overall BSM architectures defined in [1].

6.2 BSM QIDs

Central to the QoS capability of the BSM is the concept of QID's (Queue Identifiers) [3]. These represent abstract queues available at the SISAP, each with a defined class of service for transfer of IP packets to the SD layers.

The satellite dependent lower layers are responsible for assigning satellite capacity to these abstract queues according to the specified queue properties (e.g. QoS). The QID is not limited to a capacity allocation class; it relates also to forwarding behaviour with defined properties.

A QID is only required for submitting (sending) data via the SI-SAP and the QID is assigned when the associated queue is opened. An open queue is uniquely identified by the associated QID: in particular, the QID is used to label all subsequent data transfers via that queue.

The way in which QIDs are mapped to the IP layer queues is an important consideration for overall QoS.

6.2.1 QID Management

An ST QID Resource Manager (STQRM) functional entity is required to manage QIDs and their mapping BSM IDs and to the IP layer. This entity should be logically situated in both the SIAF and SDAF protocol layers, and can be part of the BSM Management Plane or Control Plane functions (ST Resource Controller).

This STQRM should also be considered as a client function to a centralised QID manager (Server) (BSMQRM) situated in the BSM QoS Manager. The BQM manages the QID resources of the BSM system and allocates QIDs and their mapping via the ST clients.

The STRC should be aware of the QID resources available at any time, either by being informed by or interrogating the BRC at IP system level (e.g. in the case of static QIDs), or by the STQRM at local SISAP level (e.g. for dynamic QIDs).

QID resources could be requested and allocated dynamically by two main paths:

1) By request through the IP layer (BRC client to BRC server by direct interaction, then forwarded to the BSM-QRM server). This latter then allocates QIDs which are distributed to the ST-QRM). The loop is closed by the BRC coordinating these resources with the NCC.

2) By request through the SD layer (BRC client to SD Resource Manager request via the control plane, followed by NCC to BSM-QRM request. This latter then allocates QIDs which are distributed to the ST-QRM). The loop is closed by the coordinating these resources with the BRC for the IP layer resources and policy.

6.3 BSM SISAP

The SISAP represents the interface between Satellite-Independent higher layer protocols and lower layer Satellite-dependent protocols.

The different cases of interaction between QoS requests and the BSM involves not only the User Plane containing the QIDs, but also the Control and Management Planes which influence the way the QIDs are used. The interaction between the IP layer QoS and the SD layer QoS takes place across the SISAP and is thus the major issue for the BSM.

The QoS functions of the User, Control and Management Planes at the SISAP are described in clause 6.5.

It is worth noting that the following U- and C-Plane functions appropriate to QoS have been defined [3].

Table 6.1: U-plane functions

Function	Description
Data Transfer	A function for sending and receiving user data via the SI-SAP. This function can be used for both unicast and multicast data transfer.

Table 6.2: C-plane functions

Function	Description
Resource Reservation	A function to open, modify and close queues in the SD layers for use by the SI layers.
Flow Control	A function to activate and configure the SD layers to provide flow control on one or more specific flows of data.

No M-plane functions are currently defined.

6.4 QoS Cases

Since QoS can be implemented in a network in different ways with different levels of complexity, several cases can be defined which illustrate typical approaches.

In each of these cases, it is assumed that the BSM system provides different levels of bearer QoS through a certain number of QID's which determine the nature of the QoS offered at the SISAP. It is the way in which the QIDs are accessed by or modified by request from the IP layer and above that changes between cases.

Starting with the simplest, these cases are described below.

6.4.1 Case 1

In this scenario there is no request or signalling of resources to or from the lower layers across the SISAP. The BSM QID resources are pre-defined, or they may at most offer varying availability for best-effort services.

There is also no admission control foreseen on a session basis. A consequence of this assumption is that no guaranteed services (e.g. IntServ GS) can be offered. This case is suited to the simplest DiffServ model where QoS is provided on a trial and error basis of resource availability.

If a several classes of service are made available via QIDs, then premium classes risk being underused by user traffic. If there is excess traffic, then in the case of DiffServ it can be re-marked or dropped according to policy.

Traffic conditioners can control the usage of the resources through enforcement of TCAs. Although a range of services can be deployed using only static marking policies, functions such as policing, shaping, and dynamic re-marking still enable services with quantitative performance metrics.

This case may include a guaranteed rate service at the SISAP (QIDs). This may be allocated to Diffserv EF traffic.

The Resource Controllers (STRC, BRC) are as simple as possible, and perform mainly policy control of DiffServ.

There is still a capability to negotiate the SLA with the customer via a Bandwidth Broker.

At the Hub side, there also a capability to negotiate services towards the network via a Bandwidth Broker.

6.4.1.1 Key Features

See figure 6.1.

In this case IP queues are managed in the User Plane only:

- No Control Plane is active at SI-SAP.
- QIDs established by management configuration (quasi-static resources).

NOTE: The properties of these QIDs may be adjusted on a time scale of hours or days.

- No RSVP/NSIS protocols.
- No Session Layer protocols (such as SIP) or SCF function.
- No admission control, just IP queue management/policing.

6.4.2 Case 2

Although BSM QID resources are still static per ST in this case, the higher layers offer richer functionality. There are session control and reservation functions, and intelligent control of IP resources by admission control in order to guarantee reserved resources. Admission control can also be used for Diffserv services.

The BRC at the Server manages the overall System resources at IP level across the system with the STRC as client. Session requests arriving at the BSM from Application Functions outside or inside the BSM (e.g. SCF) or from QoS triggers such as SIP, RSVP, NSIS, etc.) are passed to the BSM Resource Controller.

This case may include a guaranteed rate service at the SISAP (QIDs). For use of this by Intserv, IP services could be reserved up to their respective class resource limit (fixed). IP admission control is necessary to guarantee service and to avoid exceeding the resource limit. RSVP or NSIS signalling will in this case be needed to reserve resources.

A specific primitive passed across the SISAP may be used for BSM internal signalling of these session resource requests and responses (can be a generic primitive) in order to be allocated higher priority to the IP resource signalling over other SISAP primitives (such as for data).

Flow control in the C-Plane may be an option in this case (though care must be taken that flow control does not affect scheduling at the IP layer and the PHBs). Static flow control is envisaged here, which means that the activation of flow control on specific QIDS is predefined.

6.4.2.1 Key Features

See figure 6.2.

Compared with Case 1, this case adds admission control per session from external or internal BSM application layer or session layer functions SCF, SIP etc. or alternatives) and manages IP resources whist keeping the BSM resources (QIDs) static:

- More sophisticated QoS resource control in STRC and in BRC:
  - session admission control;
  - requires new BSM external interface for higher layer session-originated triggers.

BSM impacts:

- minimum C-plane at SI-SAP (contains only flow control);
- Static QIDs in BSM (resources pre-configured);
- No Resource Reservation.

C-plane primitives used: Flow Control.

6.4.3 Case 3

Compared to Case 2, this case adds dynamic resource reservation of QIDs triggered, via the BRC, by either external application triggered request (from SCF) through the BRC, by session layer QoS request or through IP QoS signalling.

Dynamic reservation of QIDs can be foreseen in two levels of complexity (representing two sub-cases):

a) modification of limited QID parameters and queue mapping only (e.g. capacity partitioning - resource request passed to the SIAF QID Resource Manager only;

b) negotiation of complete range of QID parameters (resource request passed to ST QRM and then to SD layers).

6.4.3.1 Key Features

See figure 6.3.

This case adds dynamic resource reservation to the BSM system through dynamic QIDs.

BSM impacts:

- Dynamic QIDs in BSM

C-plane primitives used below IP layer:

Case 3a:	Flow Control. Resource negotiation with STQRM only.
Case 3b:	Flow Control. Resource Reservation across SISAP.

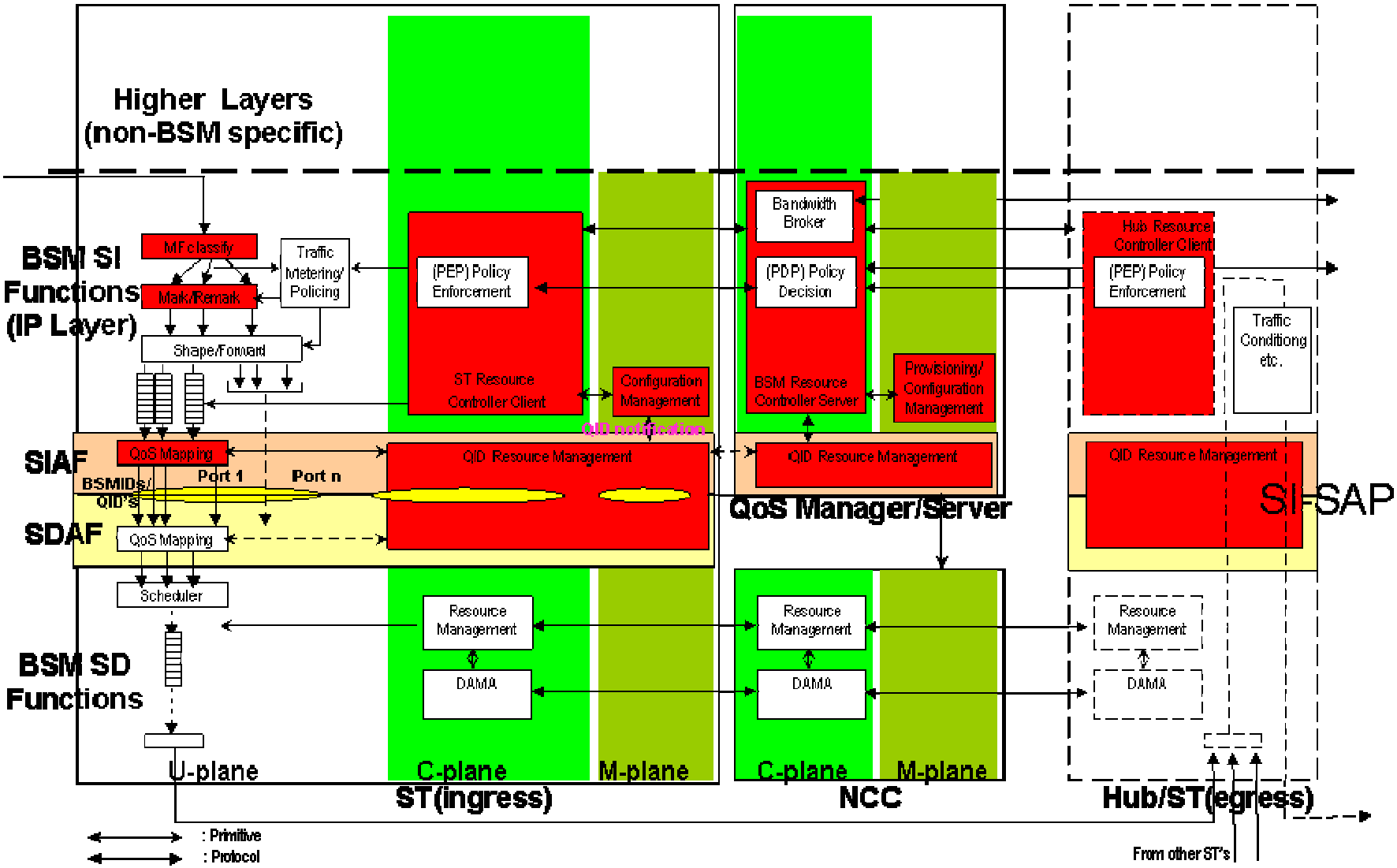


Figure 6.1: QoS Case 1

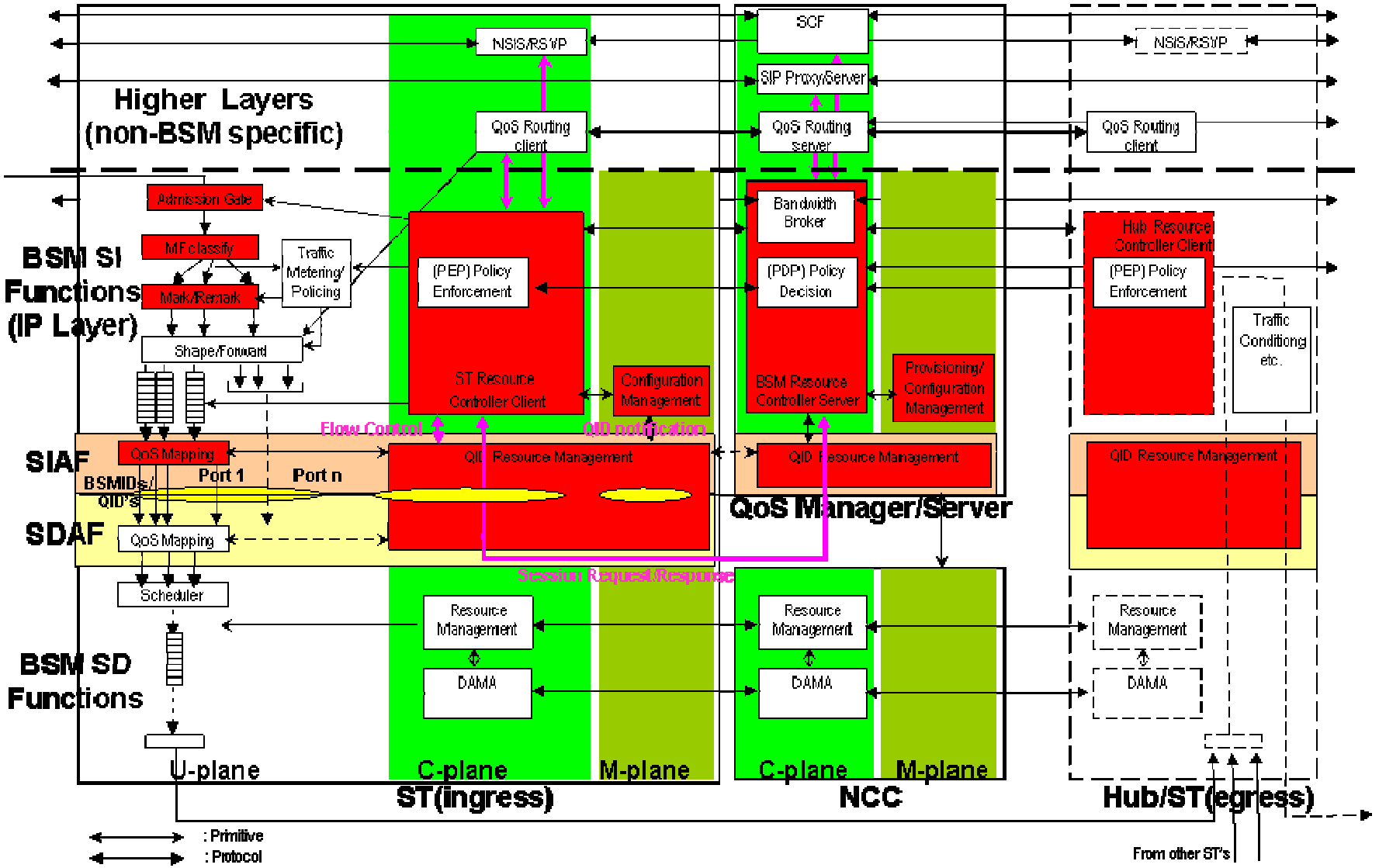


Figure 6.2: QoS Case 2

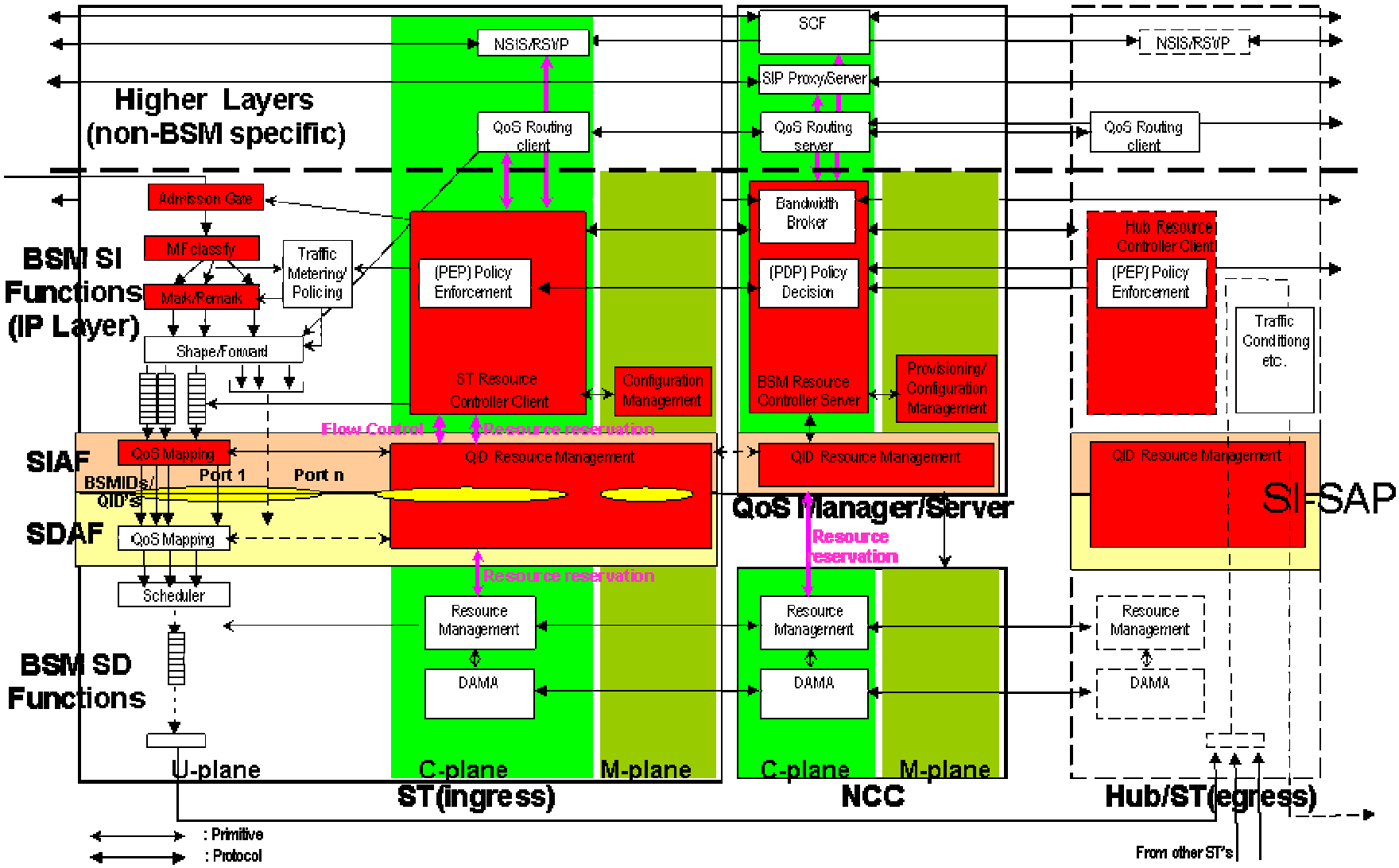


Figure 6.3: QoS Case 3

6.4.4 Functional entities

The boxes shown above in figures 6.1 to 6.3 in red indicate functional entities with at least some properties that are specific to the BSM and need further definition. Other entities are considered to be well-known standard functions.

6.4.5 Interfaces

Interfaces shown above in figures 6.1 to 6.3 are of two main types:

1) Vertical lines indicate the passing of primitives between protocol layers.

2) For ease of illustration, horizontal lines indicate peer-to-peer communication between protocol layers in different logical entities, although in reality these interfaces use data paths in the user plane protocol stack (unless otherwise indicated).

6.5 Detailed architecture at the SISAP

For simplicity, this clause describes the architecture for the most complete set of functions of the cases described above in clause 6.4. Simpler cases can be derived by omitting functions that are not required.

6.5.1 User Plane

In mesh networks the ST is the sole type of border node to the BSM and it must provide all required QoS functions. The asymmetric nature of star networks on the other hand means there are two types of border node: the single Hub station providing all ingress QoS processing from the external network, and its associated ST's.

6.5.1.1 ST Architecture

In the User Plane the ST must implement all traffic processing and policing functions according to a packet's class of service.

An ST QoS architecture is illustrated in figure 6.4.

For clarity in this figure it is assumed that there is one output port (no routing, one set of QIDs) and one subscriber per ST, either an individual or an organisation, so that one SLA and one set of IP queues is sufficient for illustration, instead of a set for each subscriber.

The ST architecture we are concerned with is based on two layers of queues: one at the IP layer and another at the SISAP identified by QIDs (further queues may exist at lower or higher layers).

The relationship between these sets of queues at the two layers is an important consideration.

Queuing could be managed solely at the IP layer, especially for relative QoS, where policies implemented would depend on available rate from the SISAP. If there are, however, different classes of service available at the SISAP, such as guaranteed rate and on-demand rate, then it would be beneficial, especially for guaranteed QoS, to map IP queues to the corresponding QIDs, in order to ensure that IP QoS parameters (such as guaranteed rate) are maintained at lower layers. The behaviour of the concatenated queues must be carefully taken into account in the overall QoS behaviour; the IP queue behaviour should not be degraded by QIDs (e.g. by increased jitter, delay, loss, etc.), and QIDs should ideally offer a transparent medium.

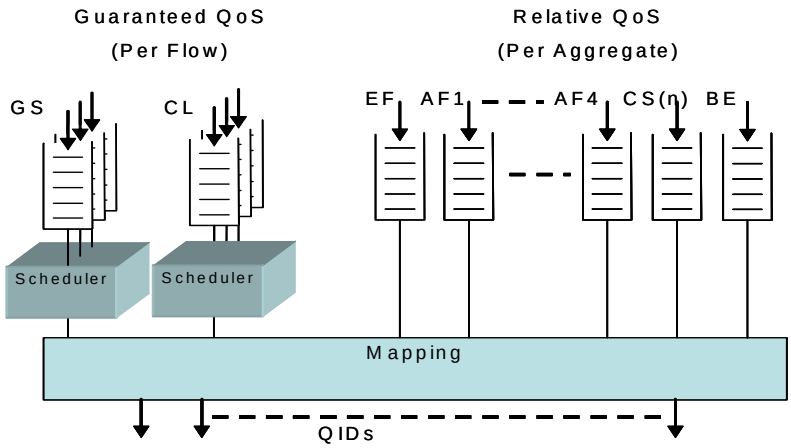


Figure 6.4: ST Queue architecture

6.5.1.2 IP Queues

6.5.1.2.1 Relative QoS

The ST queuing model shall support a typical DiffServ IP queue set for the following Per Hop Behaviours:

- Default (Best Effort - BE, defined in RFC 2474 [4]);
- Class Selector (CS - [4]);
- Assured Forwarding (AF - [5]);
- Expedited Forwarding (EF - (RFC 2598 [6]).

Other queues may be implemented depending on the PHB's defined for the BSM. A DSCP value, allocated to packets by the classifier at the ST ingress, corresponds to each queue. The queues handle aggregate IP flows of the same behaviour and control of the aggregates is typically stateless, though at the network edge admission control may be implemented.

Managing the queues includes setting the drop precedence and handling packet dropping, according to a set of rules established for each IP queue, consistent with and as an enforcement of the Traffic Conditioning Agreement (TCA) established between subscriber and the service provider (SP) as part of the SLA. This SLA includes in general support for all DSCPs. It also contains the details of the TCA for each DSCP. The TCA contains information on how metering, marking, discarding and shaping of packets must be done in order to fulfil the SLA. The TCA/SLA need to be configured in the ST. This can be done statically or dynamically; COPS protocols could be used for dynamic configuration, as per the QoS overall architecture, but other protocols (such as SNMP) may be considered as interim solutions.

BE
For Best Effort traffic, the packets are directed to a BE FIFO queue, but without any conditioning. They remain in the queue until layer 2 resources are made available (as a result of layer 2 scheduling); this does not rely on any configured bandwidth, but only on dynamic requests and contention resolution in the layer 2 scheduler. When the queue is full, a Drop Tail (DT) mechanism may be implemented to drop the packets. More sophisticated Active Queue Management (AQM) may instead be used to ensure better congestion behaviour in these conditions.

For BE there is no committed rate to adjust to, therefore no rate adjustment.

Flow control may be provided in this case from the SISAP to manage data transfer to QIDs.

CS
These PHBs ensure that DS-compliant nodes can co-exist with IP-Precedence aware nodes (with the exception of the DTS bits 3, 4, 5 - the original type of service subfield (see RFC 1349 in Bibliography). These PHBs retain almost the same forwarding behaviour as nodes that implement IP-Precedence based classification & forwarding.

EXAMPLE: As an example, packets with a DSCP value of '110000' (IP-Precedence 110) have a preferential forwarding treatment (scheduling, queuing, etc.) as compared to packets with a DSCP value of `100000' (IP-Precedence 100).

AF
Packets in each of the four AF classes may, after classification, be marked with one of the dropping precedence codes (as a function of their degree of non-conformance), before being directed to an AF queue. **Each AF queue thus needs to be actively managed** and the FIFO discipline is no longer appropriate, as it cannot ensure the required variable loss probability for AF PHB's with different drop precedence. Class Based Weighted Fair Queuing allows the bandwidth to be shared among the various classes defined. Absolute bandwidth or a percentage of the interface bandwidth may be allocated to each class. Within an AF class, packets can be dropped based on the drop precedence scheme using Weighted Random Early Detection (WRED).

The pre-defined service rate, combined with rate monitoring provided by token bucket mechanisms, ensure automatic adjustment of the rates of the incoming flows to the available rate at the SISAP. Therefore for the AF traffic **there is no need for feedback (flow control) from the SISAP** to layer 3 to provide rate control, assuming that the SISAP can provide a semi-constant rate for these classes.

EF
Delay sensitive traffic such as VoIP needs to be given high priority, and Low-Latency Queuing is suitable (LLQ). To ensure that excess voice traffic does not interfere with traffic of other classes, this priority queue is policed.

6.5.1.2.2 Guaranteed QoS

For support of guaranteed QoS using the IntServ model, additional queues for the Guaranteed Service (GS) and Controlled Load (CL) classes are introduced with higher priority than the DiffServ queues.

Unlike DiffServ, each IP flow needs a separate queue to be maintained, as well as a separate set of all the necessary mechanisms for traffic policing etc. Therefore state for each flow must be maintained in the ST.

The parameters for each flow are established dynamically by RSVP (or NSIS) signalling, or may be established quasi-statically by network management.

These queues for each of these two guaranteed classes should be served in a round-robin manner in order that any bandwidth not used by one flow is automatically allocated to other flows and each of the queues are served fairly. Fair Queuing provides to each flow a guaranteed minimum share of bandwidth together with the possibility to obtain more bandwidth. A scheduler is therefore needed at the output of each queue as shown in figure 6.4

6.5.1.3 QIDs

QIDs represent abstract queues which are associated with BSM classes of service. However more QIDs than BSM traffic classes can be assigned (e.g. with different priority within a class) in order to give increased differentiation of BSM QoS.

Dimensioning the maximum SD queue size associated with a QID is an important factor in the overall QoS. The SD queue should be dimensioned to take into account any lower layer (BoD) scheduling latency. A good practice is to dimension the SD queue to accommodate the traffic accumulated in the scheduling interval at the maximum link rate or at the user-ST interface rate, whichever is higher. This is needed in order to deal with the situation immediately after a session request acceptance, and before bandwidth resources are allocated to the link after a dynamic request

In general BE SD queues can be larger than for other classes to avoid loss, as there is no committed rate and the packets can wait for longer before capacity becomes available.

6.5.1.4 Mapping between IP queues and QIDs

The mapping of IP queues to QIDs can in principle be flexible, without being constrained to a one-to-one relationship. If less than one QID is used per IP queue in aggregate (i.e. shared QIDs), then a scheduler is needed to differentiate between priority if IP queues and ensure fair access to QIDs.

For example within the AF class, a scheduler algorithm could determine the service discipline for packets in all AF queues into one or more QIDs. The schedulers required are therefore dependent on the mapping between IP queues and QIDs.

For guaranteed services the per-flow queues for each class need to be scheduled before merging into an aggregate for which a single QID could be assigned.

The number and characteristics of QIDs assigned must therefore also be interrelated to the mapping and scheduling configuration between these sets of queues.

6.5.1.5 Flow Control

Feedback (i.e. Flow Control) may be provided via the QIDs (or from the Layer 2 queues) to the IP queues/scheduler, in order to adjust the rates of packets from the IP layer to the QID rates, which may be dynamic in rate. Such feedback would provide queue synchronization / co-ordination and can be seen as a local flow control. As a minimum, the SI-SAP should provide an interface that allows the IP-layer entities to determine the state of the available space (number of cells or bytes) in the QID-related buffers.

As a result of this flow control, queuing and scheduling would be mainly done at the IP level, as the L3 queues are also partially controlled from L2.

There is no need for feedback from EF QIDs, as it is assumed that the corresponding EF traffic packet rate is limited by the applications to a value consistent with the available rate for EF services.

6.5.2 Hub Station Architecture

Compared with typical ST's, the ingress node of the Hub Station version of an ST handles the traffic destined to many other ST's, and may be connected to several ISP networks. In many respects it can be similar to an ST used in a mesh network, though on a different scale.

IP flows are classified and processed at the ingress for each port, according to the SLA and policy applying to each ISP. The way in which users' SLA's (e.g. between users and potential ISP's, or between users and the BSM network operator) are aggregated into the SLA's between the BSM network operator and the ISP's needs to be carefully considered.

Classification should be into a common set of BSM PHB's so that after processing and forwarding the flows can be merged and queued in one set of queues for each output port.

6.5.3 Control Plane

The Control Plane is associated with call and connection control.

When an IP layer resource request is received at the ST (e.g. from RSVP or NSIS signalling, etc.) the BRC client in the ST should issue a resource request to the BRC server. This message can be sent via an IP layer flow with appropriate traffic class, or possibly via a SISAP control plane primitive dedicated to IP layer signalling flows. Several mechanisms can be envisaged for subsequent actions (see also clause 6.2.1), for example:

a) The BRC could allocate IP layer resources if there are sufficient QID resources, and reply to the ST at the IP layer.

b) If there are insufficient QID resources either:

- The BRC could immediately issue a request for SD resources across the SISAP,. This could be done via a message passed directly between the BRC Server and the NCC. If the request is accepted by the NCC, the BRC and the NCC can configure resources at the ST directly at their respective layers.

- Alternatively the BRC server could reply to the BRC client to authorise it to request SD resources across its SISAP and in the case of acceptance the STRC would be informed across the ST SISAP. The BRC would be informed by the NCC.

This SISAP layer request could be of several types according to the QoS parameters of QIDs available such as:

- bit rate;
- traffic class e.g. guaranteed rate, volume-based, etc.;
- delay;
- jitter.

A QID resource request could be for new resources or for modification of existing resources. If the request is granted then the QID parameters must be passed to the BRC client, together with the updated mapping requirement.

6.5.4 Management Plane

The Management Plane is associated with QoS provisioning and configuration management. As an example, the initial configuration of QIDs and their mapping needs to be made available to the ST. The majority of communications between management logical entities are expected to be handled by means of MIBs, but management messages in some cases could also passed more directly via the IP layer (with appropriate traffic class) between entities.

It is not considered necessary to specify a dedicated QoS management message here.

6.6 DVB-RCS Example

An example of an implementation of IP layer and SD layer queues is given in annex D for a DVB-RCS system.

Annex A (informative): End-to-End Traffic Classes

ITU-T Recommendation Y.1541 (see Bibliography) specifies network performance objectives for IP based services. The scope of the Recommendation is the UNI to UNI Interfaces as described in figure 2, user equipments characteristics are not included.

Table A.1 presents an overview network QoS class definitions and network performance objectives.

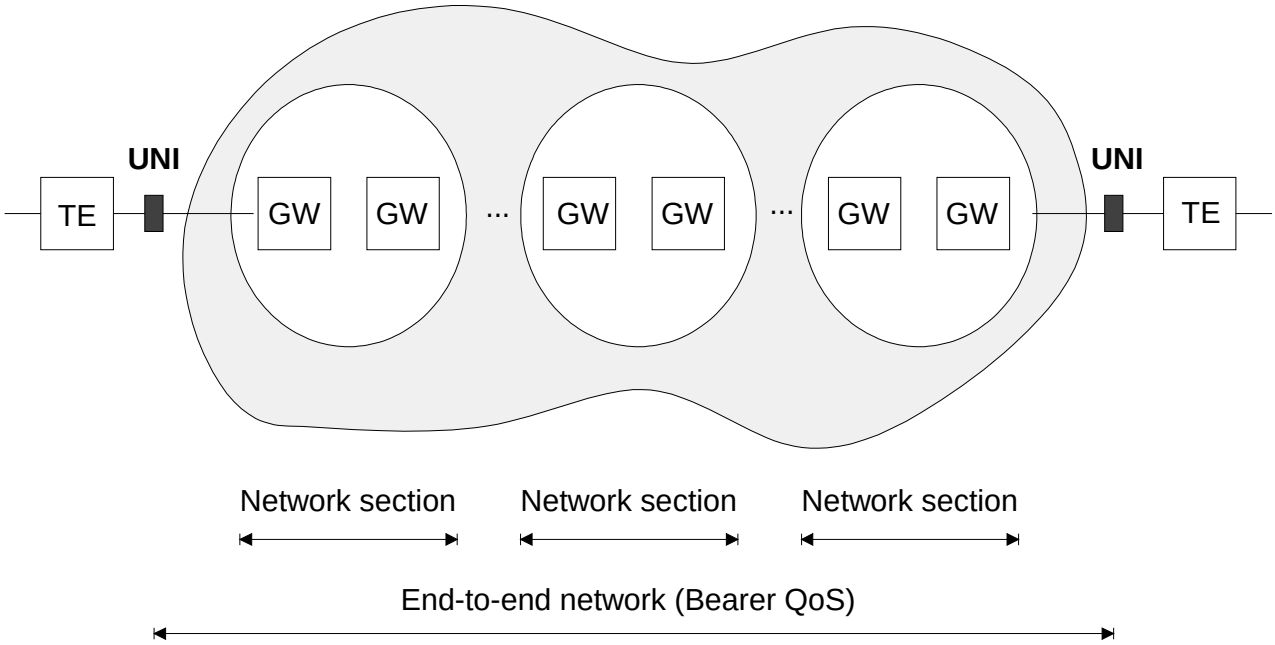


Figure A.1: Reference path for network QoS objectives defined in ITU-T Recommendtion Y.1541

NOTE: There is currently work on revision of ITU-T Recommendation Y.1541 (see Bibliography).

Table A.1: IP network QoS class definitions and network performance objectives

		QoS Classes					
Network Performance Parameter	**Nature of Network Performance Objective**	**Class 0**	**Class 1**	**Class 2**	**Class 3**	**Class 4**	**Class 5 Un-specified**
IPTD	Upper bound on the mean IPTD (Note 1)	100 ms	400 ms	100 ms	400 ms	**1 s**	U
IPDV	Upper bound on the 1-10^{-3} quantile of IPTD minus the minimum IPTD (Note 2)	50 ms (Note 3)	50 ms (Note 3)	U	U	U	U
IPLR	Upper bound on the packet loss probability	1×10^{-3} (Note 4)	1×10^{-3} (Note 4)	1×10^{-3}	1×10^{-3}	1×10^{-3}	U
IPER	Upper bound	1×10^{-4} (Note 5)					U

General Notes:
The objectives apply to public IP Networks. The objectives are believed to be achievable on common IP network implementations. The network providers' commitment to the user is to attempt to deliver packets in a way that achieves each of the applicable objectives. The vast majority of IP paths advertising conformance with Recommendation Y.1541 should meet those objectives. For some parameters, performance on shorter and/or less complex paths may be significantly better. An evaluation interval of 1 minute is provisionally suggested for IPTD, IPDV, and IPLR, and in all cases the interval must be reported.
Individual network providers may choose to offer performance commitments better than these objectives.
"U" means "unspecified" or "unbounded". When the performance relative to a particular parameter is identified as being "U" the ITU-T establishes no objective for this parameter and any default Y.1541 objective can be ignored. When the objective for a parameter is set to "U", performance with respect to that parameter may, at times, be arbitrarily poor.
All values are provisional and they need not be met by networks until they are revised (up or down) based on real operational experience.

NOTE 1: Very long propagation times will prevent low end-to-end delay objectives from being met. In these and some other circumstances, the IPTD objectives in Classes 0 and 2 will not always be achievable. Every network provider will encounter these circumstances and the range of IPTD objectives in table 1 provides achievable QoS classes as alternatives. The delay objectives of a class do not preclude a network provider from offering services with shorter delay commitments. According to the definition of IPTD in ITU-T Rec. Y.1540, packet insertion time is included in the IPTD objective. This Recommendation suggests a maximum packet information field of 1500 bytes for evaluating these objectives.

NOTE 2: The definition and nature of the IPDV objective is under study. See Appendix II for more details.

NOTE 3: This value is dependent on the capacity of inter-network links. Smaller variations are possible when all capacities are higher than primary rate (T1 or E1), or when competing packet information fields are smaller than 1500 bytes (see Appendix IV).

NOTE 4: The Class 0 and 1 objectives for IPLR are partly based on studies showing that high quality voice applications and voice codecs will be essentially unaffected by a 10.3 IPLR.

NOTE 5: This value ensures that packet loss is the dominant source of defects presented to upper layers, and is feasible with IP transport on ATM.

N.B.: The Appendices referred to in table A.1 are all appendixes to ITU-T Recommendation Y.1541 (see Bibliography).

The following example is taken from TISPAN (still draft) and similar to original BSM Traffic classes and to TIPHON classes (published):

Table A.2: TISPAN Traffic Classes

Category	Components	General QoS characteristics
Real-time conversational (e.g. telephony, teleconference, videophony and videoconference)	Speech Audio Video MM	Delay sensitive Delay variation sensitive Limited tolerance to loss / errors (depends on coding) Constant Bit Rate (CBR) and Variable Bit Rate (VBR)
Real-time streaming (e.g. audio and video broadcast, surveillance, graphics)	Audio Video MM	Tolerant to delay (buffering in terminals) Delay variation sensitive (depending on buffer sizes in terminals / gateways) Limited tolerance to loss / errors (depends on coding) Variable Bit Rate (VBR)
Near real-time interactive (e.g. web browsing)	Data	Delay sensitive (interactive services) Tolerant to delay variation Error sensitive Variable Bit Rate (VBR)
Non real-time background (e.g. Email and file transfer)	Data	Not delay sensitive Not delay variation sensitive Error sensitive Available Bit Rate (ABR)

Annex B (informative): BSM Traffic Classes

The BSM Traffic Classes are from TS 102 295 (see Bibliography) .

Table B.1: BSM Traffic Classes

BSM Traffic Class	Service Categories	Node Mechanisms	BSM Resource Management (Note 1)	Network Techniques (Informative Only)	Y.1541 class	Y.1221 Transfer Capability	PHB (Note 2)
0	Pre-emption, emergency services, essential network services	Pre-empts any traffic that has allocated BSM bandwidth	Strict admission control with pre-emption	Strict admission control with pre-emption	N/A	N/A New Traffic Class	EF
1	Real-Time, Jitter sensitive, high interaction – Fixed size cells (VoIP)	Separate queue with preferential servicing, traffic grooming, strictly admitted	Dedicated or requested bandwidth	Constrained routing and distance	0	Dedicated Bandwidth	EF
2	Real-Time, Jitter sensitive, interactive - Variable size packets (Real Time Video)	Separate queue with preferential servicing, traffic grooming, loosely admitted	Dedicated or requested bandwidth	Less constrained routing and distances	1 (with no reference to variable size packets)	Dedicated Bandwidth	EF
3	Transaction Data, Highly Interactive, (Signalling, traffic engineering, PEPs)	Separate queue, drop priority, strictly admitted	Requested or contended bandwidth	Constrained Routing and Distance	2	N/A New Traffic Class	AF
4	Transaction Data, PEP, Interactive	Separate queue, drop priority, flow controlled	Requested or contended bandwidth	Less constrained routing and distances	3	N/A New Traffic Class	AF
5	Low Loss Only (Short Transactions, Bulk Data, Video Streaming)	Long queue, drop priority, flow controlled	Requested or contended bandwidth	Any route/path	4	N/A New Traffic Class	AF
6	Medium loss, higher delay (Traditional Applications of IP Networks)	Separate queue, flow controlled	Requested or contended bandwidth	Any route/path	5	Best Effort	Default
7	Not specified Could be used for low priority broadcast/multicast traffic or storage networks (with reliable higher layer)	Separate queue	Requested or contended bandwidth	Any route/path	N/A	Best Effort	Default

NOTE 1: The BSM resource management descriptions are informative examples only and they shall not preclude a different resource management implementation.
NOTE 2: The 3 types of resource (bandwidth) allocation are indicated as dedicated, reserved and contended.
NOTE 3: Per Hop Behaviour (PHB's) are defined in RFCs 2475 and 2597.

Annex C (informative): QoS Building Block Functions

The functional blocks are allocated to the User, Control and Management Planes as described below.

Treatment of traffic is performed in two principally different ways according to whether Guaranteed or Relative QoS is applied, since Relative QoS is applied to traffic aggregates but Guaranteed QoS is applied to each flow. The differences are indicated below.

C.1 User-plane mechanisms

C.1.1 Traffic classification

Traffic classification can be done at the flow or packet level. At the ingress ST (or at the edge of the network), the function responsible for traffic classification typically looks at multiple fields (such as the five-tuples associated with an IP flow) of a packet and determines the traffic class or aggregate to which the packet belongs and the respective service level agreement.

If the flow is subject to a pre-negotiated Guaranteed QoS reservation then this flow is passed directly to the appropriate queues.

C.1.2 Packet marking

For relative QoS provision, packets can be marked according to the specific service classes that they will receive in the network.

Typically performed by an edge node (ST), packet marking involves assigning a value to a designated header field of a packet in a standard way. If done by a host, the mark should be checked and may be changed when necessary by an edge node. Sometimes, special values may be used to mark non-conformant packets, which may be dropped later due to congestion. Packets may be also promoted or demoted based on measurement results.

For example, the DS Field of the "Type of Service" byte in the IPv4 header or of the IPv6 Traffic Class byte (or the EXP bits of the MPLS shim header (RFC 3032)) is used to codify externally observable behaviours of routers in *DiffServ* RFC 2474 [4].

There is no requirement that particular DSCPs and PHBs be used for a certain service class, but as a policy it would be useful to apply them consistently across the network.

In the case of the BSM, packets may be marked according to the BSM Traffic Classes (see TR 102 295 in Bibliography).

Whether done by a host or an edge node, the criteria for packet marking need to be provisioned or configured dynamically. For dynamic configuration, the Common Open Policy Service Protocol (see RFC 2748 in Bibliography) or RSVP may be used. In the case of RSVP, the marking entity can use it to query the network about the marking to apply to packets belonging to a certain flow (see RFC 2996 in Bibliography).

C.1.3 Traffic policing

Policing deals with the determination of whether the traffic being presented is on a hop-by-hop basis compliant with pre-negotiated policies or contracts. Typically non-conformant packets are dropped. The senders may be notified of the dropped packets and causes determined and future compliance enforced by SLAs.

C.1.4 Traffic shaping

Traffic shaping deals with controlling the rate and volume of traffic entering the network. The entity responsible for traffic shaping buffers non-conformant packets until it brings the respective aggregate in compliance with the traffic. The resulted traffic thus is not as bursty as the original and is more predictable. Shaping often needs to be performed between the egress and ingress nodes.

There are two key methods for traffic shaping: leaky bucket and token bucket. The leaky bucket method employs a leaky bucket to regulate the rate of the traffic leaving a node. Regardless of the rate of the inflow, the leaky bucket keeps the outflow at a constant rate. Any excessive packets overflowing the bucket are discarded. Two parameters are characteristic to this method and usually user configurable: the size of the bucket and the transmission rate.

The token bucket method, on the other hand, is not as rigid in regulating the rate of the traffic leaving a node. It allows packets to go out as fast as they come in provided that there are enough *tokens.* Tokens are generated at a certain rate and deposited into the token bucket till it is full. At the expense of a token, certain volume of traffic (i.e. a certain number of bytes) is allowed to leave the node. No packets can be transmitted if there are no tokens in the bucket. Yet multiple tokens can be consumed at once to allow bursts to go through. This method, unlike the leaky bucket method, does not have a discard policy. It leaves to the buffer management to deal with the packets if the bucket fills up. Two parameters are characteristic to the token bucket method and usually user configurable: the size of the token bucket and the rate of token generation.

The leaky and token bucket methods can be used together. In particular, traffic can be shaped first with the token bucket method and then the leaky bucket method to remove the unwanted busts. Two token buckets can also be used in tandem.

C.1.4.1 Congestion avoidance

Congestion in a network occurs when the traffic exceeds or nears what the network can handle because of lack of resources such as link bandwidth and buffer space. A sign of congestion, for example, is that the router (or switch) queues are always full and routers start dropping packets. Packet dropping induces retransmission, which results in more traffic and worsens congestion. The chain reaction could grind the network to a halt with zero throughput. Intuition suggests very large buffers to avoid congestion owing to a shortage of buffer space. Nagle (see IEEE Transactions on communications, Vol. 35, issue 4, p 435-438 (April 1987): "On Packet Switches with Infinite Storage" in Bibliography) however showed the opposite. The long queuing delay of packets due to large buffers causes the packets to be retransmitted, which then creates congestion. Congestion avoidance deals with more robust means for keeping the load of the network under its capacity such that it can operate at an acceptable performance level, not experiencing congestion collapse.

A typical congestion avoidance scheme acts by sender's reducing the amount of traffic entering the network upon an indication that network congestion is occurring (or about to occur) ACM SIGCOMM Computer Communication Review, Symposium proceedings on Communications architectures and protocols SIGCOMM'88, Vol. 18, issue 4, p 314-329 (August 1988): "Congestion Avoidance and Control" (see Bibliography). Unless there is an explicit indication, packet loss or timer expiration is normally regarded as an implicit indication of network congestion. How the traffic source throttles back depends on the specifics of the transport protocols. In a window-based protocol such as TCP, this is done by decreasing multiplicatively the size of the window.

Ideally the source of the traffic reduction comes from a customer whose admission control priority is not critical. This may permit higher priority traffic to continue to receive normal service.

When congestion subsides, a sender then cautiously ramps up the traffic.

To avoid the potential for excessive delays due to retransmissions after packet losses, explicit congestion notification (ECN) schemes have been recently developed. RFC 3168 (see Bibliography) specifies an ECN scheme for IP and TCP among other active buffer management schemes. With the scheme, incipient network congestion is indicated through marking packets rather than dropping them. Upon the receipt of a congestion-experienced packet, an ECN-capable host responds essentially the same way as to a dropped packet.

Queue or buffer management deals with which packets, awaiting transmission, to store or drop. An important goal of queue management is to minimize the steady-state queue size while not under-utilizing link as well as avoiding the lock-out phenomenon where a single connection or flow monopolizes the queue space (see RFC 2309 in Bibliography). Schemes for queue management differ mainly in the criteria for dropping packets and what packets drop. The use of multiple queues introduces further variation in the schemes, for example, in the way packets are distributed among the queues.

C.1.5 Queuing and scheduling

In a nutshell, this mechanism controls which packets to select for transmission on an outgoing link. Incoming traffic is held in a queuing system, which is made of, typically, multiple queues and a scheduler. Governing the queuing system is the queuing and scheduling discipline it employs. There are several key approaches:

- First-in, first-out queuing: packets are placed into a single queue and served in the same order as they arrive in the queue.
- Fair queuing: packets are classified into flows and assigned to queues dedicated to respective flows. Queues are then serviced in round robin. Empty queues are skipped. Fair queuing is also referred to as per-flow or flow-based queuing.
- Priority queuing: packets are first classified and then placed into different priority queues. Packets are scheduled from the head of a given queue only if all queues of higher priority are empty. Within each of the priority queues, packets are scheduled in first-in, first-out order.
- Weighted fair queuing: packets are classified into flows and assigned to queues dedicated to respective flows. A queue is assigned a percentage of output bandwidth according to the bandwidth need of the corresponding flow. By distinguishing variable-length packets, this approach also prevents flows with larger packets from being allocated more bandwidth than those with smaller packets.
- Class-based queuing: packets are classified into various service classes and then assigned to queues assigned to the service classes, respectively. Each queue can be assigned a different percentage of the output bandwidth and is serviced in round robin. Empty queues are skipped.

C.1.6 Queue (or buffer) management

A common criterion for dropping packets is the queue reaching the maximum size. Packets are dropped when the queue is full. What packets drop depend on the drop disciplines, for example:

- "Tail drop" rejects the newly arriving packet. This is the most common strategy.
- "Front drop" keeps the newly arriving packet at the expense of the packet at the front of the queue.
- "Random drop" keeps the newly arriving packet at the expense of a randomly-selected packet from the queue. This scheme can be expensive since it requires a walk through the queue.

A scheme of dropping packets only when the queue is full tends to keep the queue in the full state for a relatively long period of time, which can have a catastrophic result in case of bursty traffic. There are schemes using a more dynamic criterion not based on the fixed maximum size of the queue and thus capable of performing active queue management. A prominent one is Random Early Detection (RED) (see IEEE/ACM Transactions on Networking, Vol. 1, issue 4, pp. 397-413 (August 1993): "Random Early Detection Gateways for Congestion Avoidance" in Bibliography), which also helps address the full queue problem and avoid congestion. RED drops (incoming) packets probabilistically based on an estimated average queue size. The probability for dropping increases as the estimated average queue size grows. In other words, if the queue has been mostly empty in the recent past, incoming packets tend to be kept; if the queue has been mostly relatively full recently, however, incoming packets are likely to be dropped. More specifically, RED employs two thresholds for the average queue size. One specifies the average queue size below which no packets are dropped; the other specifies the average queue size above which all packets are dropped. For a queue of an average size between the two thresholds, the packet dropping probability is proportional to the average size. Naturally the effectiveness of RED depends on how the relevant parameters are set. There is no single set of parameters that work well for all traffic types and congestion scenarios.

Thus appear RED variants, for example:

- Flow RED (FRED) [Lin *et al*., 1997], which introduces additional control to RED by providing differential drop treatment to flows based on their buffer usage. If the number of packets from a flow in the queue is lower than a flow-specific threshold, a newly arriving packet of the same flow will not be dropped. Otherwise, it is subject to drop treatment favouring flows with fewer packets in the buffer. Compared with RED, FRED is more flexible in protecting flows from using less- or more-than-fair share of buffer space and link bandwidth.

- Weighted RED, which introduces additional control to RED by providing differential drop treatment to packets based on their priority. The higher the priority of a packet is, the lower the probability it is to be dropped.

C.2 Control-plane mechanisms

C.2.1 Admission control

This mechanism controls the traffic to be admitted into the network. Normally the admission criteria are policy driven (see RFC 2753 in Bibliography). Whether traffic is admitted depends on an *a priori* service level agreement. In addition, the decision can depend on if adequate network resources are available so that newly admitted traffic does not overload the network and degrade service to ongoing traffic. For a service provider, maximal traffic should be admitted while the same level of QoS (including transaction performance as well as service reliability/availability expectations) is maintained for the existing traffic.

Call admission approaches related to transaction performance are typically parameter or measurement based.

1) The parameter-based approach derives the worst-case bounds for a set of metrics (e.g. packet loss, delay and jitter) from traffic parameters and is appropriate for providing *hard* QoS for real-time services. This approach is typically exercised over a resource reservation request for securing necessary resource for an ensuing traffic flow. Appendix I provides an example QoS approach making use of such a type of admission control.

2) In contrast, the measurement-based approach uses measurements of existing traffic for making an admission decision. It does not warrant throughput or hard bounds on packet loss, delay or jitter and is appropriate for providing *soft* or relative QoS. This approach has in general higher network resource utilization than the parameter-based one. Note that in principle it is possible to have a hybrid approach such as using measurements to update the resources available in the parametric approach.

Admission control can also be used to meet requirements for service reliability/availability over a specified period for the desired transaction types as negotiated in the SLA. Specifically, the desired service reliability/availability can be requested as a priority level for admission control that, in turn, determines the setup of a "connection" or link such as an LSP. Admission control policies give preference to traffic streams (e.g. for emergency communications) deemed to be more critical by a service provider under conditions of congestion. Admission control priority is a way of giving preference to admit higher priority LSPs ahead of lower priority LSPs.

Annex A further specifies the priority levels for admission control.

C.2.2 Resource reservation

This mechanism sets aside required network resources on demand for delivering desired network performance. Whether a reservation request is granted is closely tied to admission control. All the considerations for admission control therefore apply. But in general a necessary condition for granting a reservation request is that the network has sufficient resources.

The exact nature of a resource reservation depends on network performance requirements and the specific network approach to satisfying them. For example, in the *IntServ* approach, simplex flows are what matter and are characterized in terms of parameters describing a token bucket, and receiver-initiated reservations are done on demand according to peak rate requirements to guarantee delay bounds. Regardless of the specifics, it is important for service providers to be able to charge for the use of reserved resources. Therefore, resource reservation needs support of authentication, authorization, and accounting and settlement between different service providers. Resource reservation is typically done with a purpose-designed protocol such as RSVP (see RFC 2205 in Bibliography).

Resource reservation can be thought of as a distributed or a centralized functionality. The discrepancy of actual versus the predicted resource availability is a major issue and care should be given to use the most current information, making the node, link and other resources available for the requesting application.

C.3 Management-plane mechanisms

C.3.1 Policy

Policies are a set of rules typically for administering, managing and controlling access to network resources. They can be specific to the needs of the service provider or reflect the agreement between the customer and service provider, which may include reliability and availability requirements over a period of time and other QoS requirements. Service providers can implement mechanisms in the control and data planes based on policies. Some potential applications are policy routing (directing packet flow to a destination port without a routing table), packet filtering policies (marking or dropping packets based on a classifier policy), packet logging (allowing users to log specified packet flows) and security-related policies.

Various events can trigger policy decisions. Some are traffic related and some are not. The details usually depend on specifics of the applications. RFC 2748 (see Bibliography), for example, specifies a simple query and response protocol that can be used to exchange policy information between a policy server (or policy decision point) and its client (or policy enforcement point).

C.3.2 QoS routing

In its narrow definition, QoS routing concerns the selection of a path satisfying the QoS requirements of a flow. The path selected is most likely not the traditional shortest path. Depending on the specifics and the number of QoS metrics involved, computation required for path selection can become prohibitively expensive as the network size grows. Hence practical QoS routing schemes consider mainly cases for a single QoS metric (e.g. bandwidth or delay) or for dual QoS metrics (e.g. cost-delay, cost-bandwidth, and bandwidth-delay).

NOTE: Note that some of these metrics are additive and some of them are limiting. For example, delay and cost are additive, bandwidth is limiting. These considerations are important in devising implementable routing algorithms.

To further reduce the complexity of path computation, various routing strategies exist. According to how the state information is maintained and how the search of feasible paths is carried out, there are strategies such as source routing, distributed routing, and hierarchical routing (see IEEE Network, Special Issue on Transmission and Distribution of Digital Video: "An Overview of Quality-of-Service Routing for the Next Generation High-Speed Networks: Problems and Solutions" in Bibliography). In addition, according to how multiple QoS metrics are handled, there are strategies such as metric ordering and sequential filtering, which may trade global optimality with reduced computational complexity (see RFC 2386 in Bibliography).

The path selection process involves the knowledge of the flow's QoS requirements and characteristics and (frequently changing) information on the availability of network resources (expressed in terms of standard metrics such as available bandwidth and delay). The knowledge is typically obtained and distributed with the aid of signalling protocols. For example, RSVP (see RFC 2205 in Bibliography) can be used for conveying the flow's requirements and characteristics and OSPF extensions as defined in RFC 2676 (see Bibliography) for resource availability. Compared with shortest-path routing that selects optimal routes based on a relatively constant metric (i.e. hop count or cost), QoS routing tends to entail more frequent and complex path computation and more signalling traffic.

It is important to note that QoS routing provides a means to determine only a path that can likely accommodate the requested performance. To guarantee performance on a selected path, QoS routing needs to be used in conjunction with resource reservation to reserve necessary network resources along the path.

QoS routing can also be generalized to apply to traffic engineering. (Concerning slowly-changing traffic patterns over a long time-scale and a coarse granularity of traffic flows, traffic engineering encompasses traffic management, capacity management, traffic measurement and modelling, network modelling, and performance analysis.) To this end, routing selection often takes into account a variety of constraints such as traffic attributes, network constraints, and policy constraints (see RFC 3272 in Bibliography). Such generalized QoS routing is also called constraint-based routing, which can afford path selection to bypass congested spots (or to share load) and improve the overall network utilization as well as automate enforcement of traffic engineering policies.

The ITU-T Recommendation E.360.x series of Recommendations (see Bibliography) describe, analyse, and recommend methods for controlling a network's response to traffic demands and other stimuli, such as link or node failures. Specifically, the methods addressed in the E.360.x series include call and connection routing, QoS resource management, routing table management, dynamic transport routing, capacity management, and operational requirements. ITU-T Recommendation E.361 (see Bibliography) further specifies QoS routing functions and associated parameters, such as bandwidth allocation and protection, routing priority, queuing priority, and class-of-service identification. In addition, ITU-T Recommendation E.361 prescribes means for signalling QoS routing parameters across networks employing different routing technologies.

C.3.3 Service level agreement

A Service Level Agreement (SLA) typically represents the agreement between a customer and a provider of a service that specifies the level of availability, serviceability, performance, operation or other attributes of the service. It may include aspects such as pricing that are of business nature. The technical part of the agreement is called the Service Level Specification (SLS) (see RFC 3198 in Bibliography), which specifically includes a set of parameters and their values that together define the service offered to a customer's traffic by a network. SLS parameters may be general such as those defined in ITU-T Recommendation Y.1540 or technology specific such as the performance and traffic parameters used in *IntServ* or *DiffServ*. Overall, ITU-T Recommendation E.860 (see Bibliography) defines a general SLA framework for a multi-vendor environment.

C.3.4 Provisioning

C.3.5 Billing (Traffic metering and recording)

Metering concerns monitoring the temporal properties (e.g. rate) of a traffic stream against the agreed traffic profile. It involves observing traffic characteristics at a given network point and collecting and storing the traffic information for analysis and further action. Depending on the conformance level, a meter can invoke necessary treatment (e.g. dropping or shaping) for the packet stream.

Annex D (informative): Example of DVB-RCS Queue implementation

D.1 Layer 3 QoS mechanisms

D.1.1 Layer 3 Qos Support On Forward Link

The gateway is the entry point into the satellite network DS domain (which coincides with the INAP domain). The edge router should therefore behave as a boundary node and implement policy / traffic conditioning functions, in addition to packet classification and per hop forwarding according to packet's class of service. The edge router is part of the TISS in the access reference architecture; its functionality as a boundary node is illustrated in figure D.1.

NOTE: FLSS can also behave as an boundary (edge) node, if the TISS does not implement QoS functions.

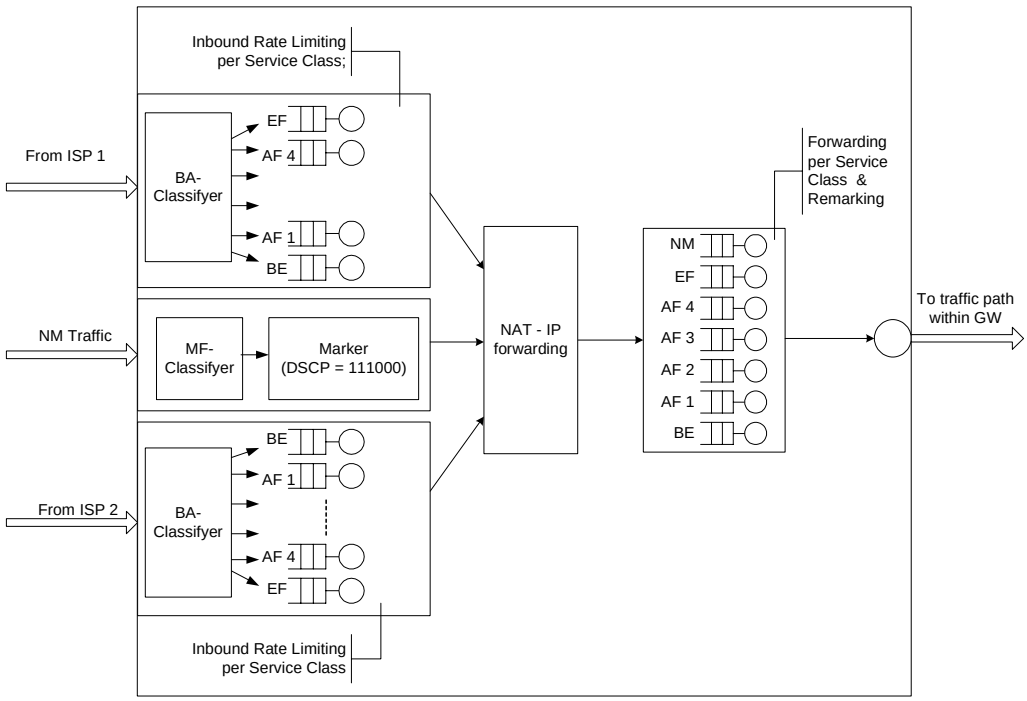


Figure D.1: The edge router as a Diffserv boundary node

Two ISPs have been illustrated in figure D.1, ISP1 and ISP2. The assumption was made that the incoming packets have been marked with the appropriate DSCP at the ISP, therefore simple BA classifiers can be implemented; they need to be configured with the rate limiting functions (for the purpose of policing) specific to the service classes supported by the INAP. The policing rules should be consistent with the SLA defined for given service class. By contrast, the Network Management (NM) packets need to be marked (by a Multi-Field (MF) classifier) to an agreed service level and then rate-limited (shaped), based on filtering rules that are part of the ISP-INAP SLA. NM packets correspond to either local management messages (i.e. OA&M) or to inter-network messages (associated with session/QoS signalling, for example).

After being policed, the packets reach the NAT-IP Forwarding block, which acts primarily as an IP multiplexer: it multiplexes traffic from different ISPs and then routes them to a local traffic path (i.e. an FLSS within the Gateway) for transmission over the air interface. The routing is via QoS-specific processing blocks (i.e. sets of queues), which perform classification and scheduling functions (for service differentiation), consistent with CoS precedence levels.

Within a given gateway there are several local traffic paths, each relying on a number of IP components (e.g. PEP, IPSec Server, IP Switch, all part of the TISS) and including an FLSS. The assumption is made that the packets will be routed to one forward path or another, based on their IP addresses. All IP components on a given forward path should be regarded as DS interior nodes. As such they have to perform packet classification and class-specific forwarding only (i.e. no policing). Alternatively, all packets from a given ISP could be directly forwarded to an FLSS, after multiplexed with the management traffic. This will depend on the TISS functional architecture.

Within the FLSS the IP/DVB Gateway is the only component that is QoS-aware. This is because QoS is not propagated within the MAC (MPEG) layer: after conversion, the MPEG packets are forwarded via virtual MPEG pipes, without changing their order within the pipes (see clause D.2.3). These virtual pipes correspond to the ISP-INAP SLA model.

The QoS functional block of the IP/DVB Gateway is represented in figure D.2.

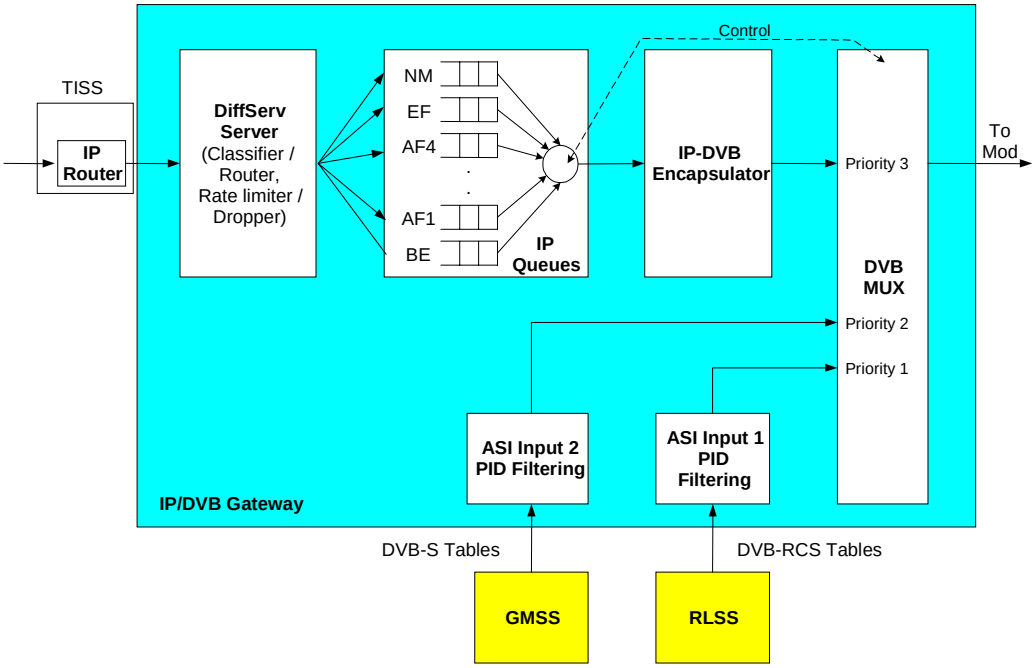


Figure D.2: IP/DVB Gateway functional QoS architecture

The IP part of the IP/DVB Gateway has the attributes of an interior node. Separate queues are implemented for each supported DS class and the packets from these queues are scheduled based on a pre-defined discipline (scheduling algorithm). After encapsulation, the user traffic is multiplexed with signalling traffic (DVB tables), either from the GMSS (standard tables) or from the RLSS (RCS tables). The local multiplexer (DVB MUX) also performs a scheduling function, giving the highest priority to the DVB tables (priority 1 and 2) and the lowest priority (priority 3) to the user packets. Please note that at this stage the user packets have already been ordered (in the stream) based on DS precedence and this order will be preserved all along the MPEG processing path.

NOTE: This will not be true for the DVB-S2 standard.

As a result of multiplexing user traffic and signalling the capacity available for user traffic is not fixed: it varies function of the amount of signalling (which has higher priority) sent at a given time. The rate of the multiplexed traffic and signalling should not exceed the output rate (which is limited for a given TDM). To this end the total amount of traffic (IP packets) that can be scheduled from the QoS queues is modulated/controlled by the DVB MUX. An IP rate limiting function will be needed in the DiffServ Server, in addition to packet classification/forwarding. Rate limitation would require configuring class-specific profiles (max/min rates, as per SLA) in the IP router and may lead to the discarding of out-of-profile traffic, as needed - function of the load offered in each class. The IP router is part of the TISS and is used to route both FL and RL IP packets.

The above functionality is not yet available in the commercial IP/DVB Gateways. SkyStream products have started to implement QoS features, such as separate input ports with separate queues per port, that can be associated with QoS classes; more features can be added as the need arises.

The capacity (or rate) of a virtual pipe reflects the aggregated needs of all subscribers receiving traffic in a given class, regardless of the traffic originating ISP. Alternatively, virtual pipes could be set-up per ISP. The way the aggregation is performed will depend largely on the ST/Host network model and on the Subscriber SLA specifics (single/multiple user(s) per ST, single/multiple subscriber(s) per ST, SLA per user/ST, single QoS class per ST, etc.). Offering FL QoS to end-users will mainly depend on the support offered by the ISP. In terms of the Gateway design, the impact will be on how user's SLAs are aggregated into the ISP-INAP SLA.

D.1.2 Layer 3 Qos Support On Return Link

D.1.2.1 Overview

On the RL the ingress node into the DS domain (INAP domain) is the ST, while the egress point is the edge router within the Gateway/TISS. This represents a major difference in comparison with the situation on the FL, in two respects:

- The ingress point is distributed across the ST population, while on the FL it is unique.
- On the RL each ingress node only handles the ST's traffic, while on the FL the ingress node handles the traffic of all STs.

IP packets entering the INAP network via an ST are transported to the Gateway via the air interface. The RL path within the Gateway includes the RLSS and a number of IP gateway components (ATM/IP router, IPSec, PEP) in addition to the edge router (all part of the TISS). The MAC Scheduler within the RLSS implements the dynamic RL resource control, and thus will have an important role to play regarding RL QoS provisioning. With the exception of the edge router (which is the egress point and might be required to perform packet DSCP re-marking), the other Gateway IP network components are DS interior nodes and as such they only need to implement packet classification and class-specific forwarding.

D.1.2.2 ST role in QoS support on the RL

As a boundary node the ST is the single most important component regarding QoS support on the RL. It has to implement traffic conditioning/policing functions, in addition to packet classification and per hop forwarding/scheduling according to packet's class of service. A possible ST QoS architecture, based on the DiffServ framework, is illustrated in figure D.3.

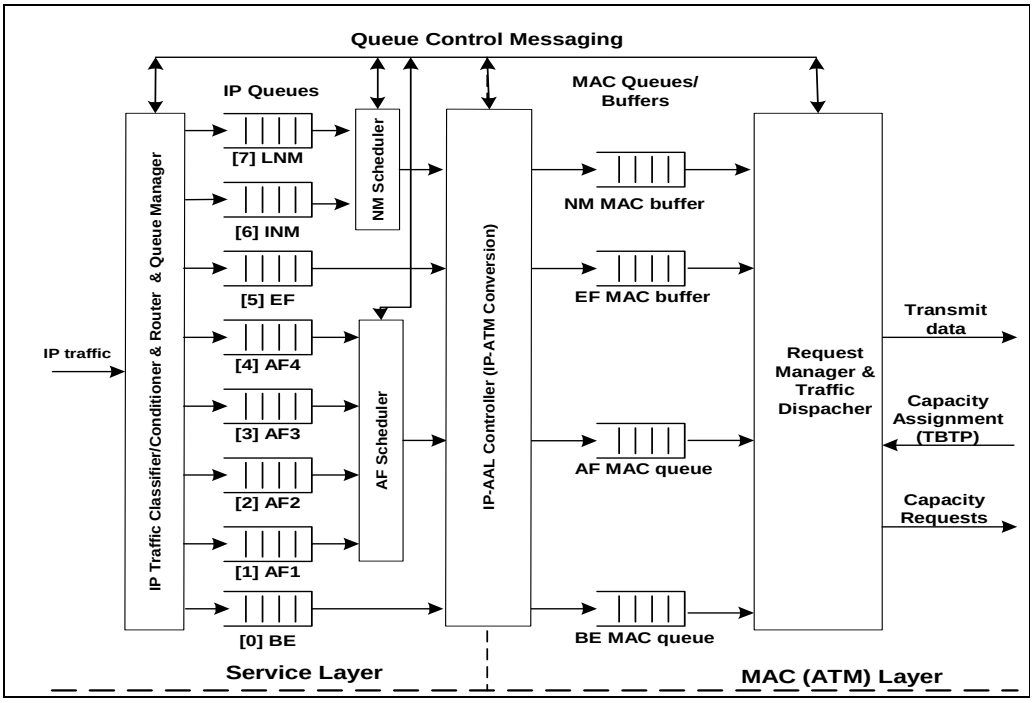


Figure D.3: ST DiffServ architecture

The architecture assumes that there is one subscriber per ST. The ST will share the RL resources in the beam with other STs/subscribers of the same ISP.

The architecture is based on two sets of queues: one in the IP domain and another one in the MAC domain (ATM for access topology, MPEG for mesh[3]). The queuing, actively managed, takes place primarily in the IP queues, while the MAC queues are rather used as buffers.

The queuing model supports a total of eight IP queues, namely the BE queue, four AF queues, an EF queue, an INM queue and an LNM queue. The figures in square brackets [ ] against each queue define the code point for each queue type and correspond to the value of the Precedence field. The BE, INM and LNM queues correspond to the original heritage traffic precedence defined in RFC 791 (see Bibliography). INM and LNM queues are reserved for management traffic. In a satellite network such traffic is currently limited to local management traffic (i.e. OA&M traffic as LNM). The design evolution towards supporting end-to-end QoS involving session and QoS signalling will probably lead to the need to support INM traffic as well, therefore the full set of queues may be needed.

The **IP Traffic Classifier/ Conditioner/ Router & Queue Manager** function is in charge of assigning the incoming IP packets to one of the eight IP queues and conditioning them accordingly. More particularly, it is responsible for:

- **Traffic Classification:** the assumption is made that the incoming packets are already marked by the applications run on the user's host, so that only a Behaviour Aggregate (BA) classifier is needed. Marking / remarking of a packet by the ST would require to implement a Multi-Field (MF) classifier in the ST (or in the host private network), which would require that the host trusts the network (for marking). MF classifier may also be needed for marking the management traffic.

- **Traffic Conditioning** (metering, shaping, dropping).

- **Managing IP Queues:** including setting the drop precedence and handling packet dropping, according to a set of rules established for each IP queue, consistent with and as an enforcement of the Traffic Conditioning Agreement (TCA) established between subscriber and its ISP (as part of the SLA).

IP Traffic Classifier / Conditioner / Router & Queue Manager functions would be implemented as Traffic Conditioning Blocks, typically one for each class of service. A TCB implements traffic conditioning based on the TCA.

The subscriber-ISP SLA includes in general support for all DSCPs. It also contains the details of the TCA for each DSCP. The TCA contains information on how metering, marking, discarding and shaping of packets must be done in order to fulfil the SLA. The TCA/SLA need to be configured in the ST. This can be done statically or dynamically; COPS protocols could be used for dynamic configuration, as per the QoS overall architecture, but other protocols (such as SNMP) may be considered as interim solutions.

Please note that traffic conditioning applies to behaviour aggregate and not to individual flows from end users. Providing QoS per user and/or application further raises the ST complexity, therefore it will not be considered in the early implementation stages. However, in the future, one can conceive per user/application QoS support (see clause D.1.2.3).

A reduced set of queues is suggested at the MAC layer, as each queue requires its own PVC and the number of PVCs used in the system may be limited by system and/or equipment constraints. In general, one queue is recommended for each basic class of service (or PHB group), i.e. EF, AF and BE. One additional queue is typically required for NM messages; alternatively NM messages could be aggregated with the user traffic in one of the traffic classes. MAC queues can be associated with MAC QoS classes of service. For more details on QoS handling at MAC layer (level 2) in terminals please refer to clause D.2.3.

As only one AF MAC queue is used, the AF scheduler is responsible for differentiating between AF classes. The AF scheduler algorithm determines the service discipline for AF packets in all IP AF queues into one MAC queue.

The interface between the two sets of queues (domains) is via AAL5 for the ATM profile illustrated in figure D.3.

The IP packets at the output of each IP queue are converted into ATM cells (by the **IP-AAL5 Controller**, using IP over ATM), and transferred to the appropriate MAC queue/buffer. Depending on packet length and available transmit bandwidth, several IP packets may be converted/stored in this manner to avoid underflow in this buffer. At the same time, feedback may be provided from the MAC queues to the IP queues / AF scheduler, in order to adjust the rates of the incoming traffic to the rates available on the return channel, which are subject to dynamic assignment; such feedback would provide queue synchronization / co-ordination and can be seen as a local flow control. As a minimum, the MAC layer should inform the IP layer about the available space (number of cells or bytes) in the AF (BE) MAC queues, in order to prevent the transfer of incomplete IP packets to the MAC queues/buffers. These buffers need to be appropriately dimensioned in order to cope with the scheduling latency. A good practice is to dimension the queues in order to accommodate the traffic accumulated in the scheduling interval at the maximum return link rate or at the user-terminal interface rate, whichever is higher. This is needed in order to cope with the situation immediately after logon, before traffic resources are allocated on the return link as a result of dynamic requests (i.e. before the first TBTP is received). In general the BE queue will be larger, as there is no committed rate and the packets can wait for longer before capacity becomes available.

There is no need for feedback from the EF MAC buffer, as it is assumed that the corresponding EF traffic packet size is limited by the application to a value consistent with the available rate for EF services. With regard to the NM traffic, a feedback from the MAC layer to IP layer may be useful or not, function of the MAC service offered to NM traffic (see Table D.1 in clause D.2.4.4.5).

As a result of the feedback from the MAC level to the IP level the queuing is mainly done at the IP level. The sizing of the AF and BE IP queues is more complex, as it depends not only on the status of the MAC queues but also on the queuing/discarding policies and scheduling algorithms.

In practical implementations there may be a single set of queues, with IP inputs and ATM outputs (i.e. AAL5 buffers are used as MAC queues).

The DiffServ mechanisms are implemented in the IP domain, while the transport of traffic over the air interface takes place in the MAC domain and is governed by the MAC layer protocol implemented in the MAC Scheduler. In order to meet the DiffServ forwarding requirements the IP classes of service need to be appropriately mapped into MAC QoS classes and then into capacity categories supported by the Scheduler. Such a mapping is suggested in clause D.2.4.4.5. The differentiation between classes will be done by setting the queue attributes, at both IP and MAC levels. At IP level the attributers are derived from the Subscriber-ISP TCA/SLA. For MAC queue configuration please refer to clause D.2.4.4.5.

The **Request Manager & Traffic Dispatcher** function handles two functions: capacity request function and assignment distribution function. Please refer to clause D.2.4.4.5 for details.

Queue Control Messaging provides for the co-ordination and synchronisation of the various functional blocks, including the feedback from the MAC queues to the IP queues (for local flow control).

D.1.2.3 QoS levels on the return link

QoS levels in this clause should be understood as relative rather than absolute and only with regard to traffic forwarding within the satellite network. True end-to-end QoS requires additional components, as identified in the overall QoS architecture.

The QoS on the return link depends primarily on the ST design and to a lesser extent on the RLSS design. The ST architecture proposed in figure D.3 is rather general, and can support IP-QoS consistent with any DSCP. However, in the early deployment stages, only a limited number of classes may be needed. Supporting a particular DSCP (or a subset of DSCPs) on the RL is only a matter of configuration of the ST and RLSS. An ST only needs to be configured with the TCB and the queue specific to the supported class, and only one PVC per ST is needed (and possibly an additional one for NM traffic).

In the ST QoS architecture only one subscriber per ST was assumed. Multiple subscribers can be supported, but this may require the duplication of the architecture in figure D.3 for each subscriber, function of the specifics of SLAs (e.g. whether the Subscriber SLAs are with the same ISP or not, whether there are soft or hard boundaries between the resource owned on the RL by different ISPs).

QoS is applicable to traffic aggregates and not to individual flows. For QoS support at flow level two approaches can be taken:

- Propagate the flows within the INAP.
- This would require virtual circuits for individual flows and the implementation of per-flow states and forwarding mechanisms in various network components. The former aspects raises issues related to the extent to which return link capacity (per ST) can be fragmented, given the relatively low rates on the RL. The latter may raise scalability issues.
- Define rules/policies for CAC implemented in STs.
- This approach is in line with the evolution of QoS support towards dynamic SLAs and policy based admission control. Per user/application QoS support can be conceived based on rules implemented in ST and a client-server relationship between ST and its hosts/users. This would be equivalent to a local CAC function, based on a PEP-PDP architecture, with the PDP function implemented in ST (Local PDP – LPDP) and the PEP implemented at host. The policies/rules, yet to be defined, would be downloaded into the LPDP from the Gateway (ACSS) / NCC.

With both approaches only a limited number of flows in the network can be provided. Adequate signalling would be required in both cases.

Supporting QoS at flow level based on the first approach can be associated with semi-permanent or on-demand connections established within INAP; such a connection can be set-up considering the needs of a particular flow. Layer 3 QoS support on mesh links.

For mesh topology the uplink coincides with the uplink of the return link and the downlink with the downlink of the forward link. The mesh processor on board the satellite maps the return link traffic into downlink traffic and for a given mesh channel the downlink capacity always matches the uplink capacity. QoS management for mesh network would therefore only requires the management of uplink resources, and will thus be similar to QoS management on the return link of the access network. The main differences are as follows:

The ST QoS architecture in figure D.3 should be duplicated for each supported channel.

The MPEG profile would require MPE instead of AAL5. As a result the connections will be identified not by PVCs but by PIDs / MAC addresses.

The mesh SLA should be negotiated with the mesh network manager and not with a public SP. It should have per-channel components.

In a pure mesh network the QoS control hierarchy would only involve two domains: the satellite network domain (INAP) and the user domain. The user domain remains unchanged. The network domain would not need a gateway, as user data traffic only flows between STs. The boundaries of the QoS domain (INAP domain) are thus defined by the STs. The NCC is hosting the MAC control function (MAC Scheduler) and is also responsible (via the ACSS) for the ST control (e.g. configuration of parameters, logical resource assignment).

The reference mesh topology relies on a regenerative satellite, which also allows access connections. In this context a gateway (GW-ST) is used and the overall QoS architecture also include a public domain (ISP domain).

D.2 Layer 2 (MAC) QoS mechanisms

D.2.1 Introduction

Level 2 QoS mechanisms refer to link layer (access layer or MAC layer) functions defined in the user plane and control plane. The functions in the user plane ensure the transport of traffic in the format(s) defined for the link layer. The functions in the control plane are responsible for the management/provisioning of satellite resources needed to meet the traffic forwarding requirements associated with the QoS classes supported by the overall system.

Two QoS management architectures have been provided: one for the satellite access topology, the other for the satellite mesh topology. The level 2 mechanisms in this clause apply to both topologies, to the corresponding links (i.e. return/forward links for access, mesh links for mesh). The differences in implementation will be highlighted where appropriate.

Satellite resources (i.e. bandwidth/capacity and the associated logical identifiers) are defined in the INAP domain. They are organised in consistency with the DVB-RCS air interface, based on a TDM scheme on the forward link and on a MF-TDMA scheme on a return link. However, the organisation can be different for the two topologies.

Resource organisation will reflect satellite interactive network connectivity. For access topology the connectivity is defined between a number of user beams and a service beam. The service beam is the beam where the access Gateway is located. The assumption is made that there is only one Gateway per INAP. For mesh topology the connectivity is defined between user beams (for traffic), but also between user beams and a service beam (for control). The service beam is the beam where the NCC is located. It may also include a Gateway (GW-ST), since the mesh topology (based on a regenerative satellite) does not preclude access services.

The net result of the above is that in mesh topology the STs within a beam can have multiple connection paths, to other STs in the same or different beams, and to a Gateway. The connection paths are associated with "fast circuits" established by the OBP, as per its switching configuration (controlled by the NMC). At the time scale considered for level 2 resource management the circuits are considered static. Changing the OBP configuration and the implications on the layer 2 mechanisms are considered outside the scope of this study.

One can say that the INAP offers connectivity services. The DVB services, required for the operation of the interactive network, rely on these connectivity services. The DVB services are then used to provide IP services, the ultimate goal of the satellite network.

In the following subclauses the level 2 mechanisms will be analysed in both user plane and control plane, with regard to logical resources and physical (capacity) resources. The analysis will highlight the differences between the access service and mesh service, between forward link and return link.

D.2.2 Layer 2 Resource Organisation

D.2.2.1 Logical resources (layer 2 addressing)

Layer 2 logical identifiers include MAC addresses, PID values, VCI/VPI (PVC), Group_id, Logon_id Population_id, Channel_id. Different identifiers are used on various links defined for the access topology and mesh topology.

D.2.2.1.1 Access topology

For access topology the layer 2 relies on ATM format on the return link (via AAL5 encapsulation) and on the MPEG2 TS format on the forward link. Consequently the following identifiers will be used:

MAC address

The MAC address is a physical address stored in STs a non-volatile memory and corresponds to a unique ST hardware identifier. It shall comply with the IEEE 802.3 standard and shall consist of 48 bits. The value 0xFFFFFFFFFFFF shall be reserved for broadcasting to all STs. MAC address will be used:

- In the forward direction, to encapsulate the IP datagrams into MPEG2-TS packets (MPE), as specified in TS 101 192 (see Bibliography).
- In the return/forward direction: in some DVB-RCS signalling (CSC, TIMs).

PID

The PID is primarily used on the forward path (i.e. it is a DVB-S identifier) for sub-net filtering of MPEG packets at the terminal. PID filtering is followed by MAC filtering (of DSM-CC sections) and then by IP filtering, so an ST can reassemble its IP datagrams.

VPI/VCI

VPI/VCI (or PVC) is used to identify virtual channels on the return links. A virtual channel is established between a ST (as source) and a destination, defined by the component in the access Gateway where the ATM to IP conversion takes place (e.g. the IP/ATM Switch). It is used at the destination to segregate the ATM cells transmitted on a given virtual channel from cells in other virtual channels, in order to enable the reassembly of IP packets.

More than one PVC can be used per ST, function of the number of implemented QoS classes of service and the corresponding MAC queues (one PVC per MAC queue). For a discussion regarding the need for multiple PVCs please refer to clause D.4.

The VPI/VCI field in the ATM cells header is 24-bit long.

DVB-RCS Identifiers

Group_id

The Group_id identifies a group of terminals receiving a common service. It is used with regard to the operation of the return link, as per the DVB-RCS standard: STs of the same Group_id will receive their assignments via the same section in the TBTP and the error messages via the same section of the CMT.

The Group_id consists of 8 bits. The value 0xFF is reserved for system use.

It is expected that an ST will remain part of the same Group_id in successive logon sessions.

Logon_id

The Logon_id identifies uniquely an ST within a Group_id for the duration of one logon session.

The Logon_id consists of 16 bits. The value 0x3FFF is reserved for system use.

Channel_id

The Channel_id is used on the uplink return link, typically to identify resources associated with a destination (i.e. a connection path). In this context it is primarily used in the management of the uplink return link resources (by the Scheduler).

The Channel_id consists of 4 bits (max 16 destinations). In the case of access topology there is only one destination : the access Gateway.

In general, the channels define partitions of the return link resources that are treated independently from resources in other channels. In this context the Channel_ids can also be used to differentiate between MAC QoS classes. This usage of Channel_id is not envisaged in the study.

Population_id

The population_id is typically used on the forward link to identify a group of terminals receiving the same Forward Link Signalling (FLS) service.

In summary, the forward link makes use of the MAC@, PID, Group_id, Logon_id, Population_id, while the return link makes use of the MAC@, Group_id, Logon_id, PVC.

D.2.2.1.2 Mesh topology

For mesh topology layer 2 relies on MPEG2-TS format on both return link and forward link. Consequently the following identifiers will be used:

MAC address

As MAC address is ST-specific, its definition is the same regardless of the network topology. However, its usage is different. In the case of mesh topology it is used on both uplink and down link (in the MPE process). In addition, it is used on the return/forward links (to/from NCC) in some DVB-RCS signalling (CSC, TIMs).

Please note that in mesh topology two STs are involved in a connection, each with its own MAC address.

PID

In the case of mesh topology the PIDs are used on both uplinks and downlinks. On downlinks they are used not only for traffic (from other mesh STs), but also for signalling (from the NCC). It is assumed that the OBP multiplexes the traffic and signalling to a given ST population in the same transport stream.

The usage of PIDs for signalling (DVB tables) on downlinks is similar to their usage on forward links in the access topology, while the usage for traffic is different.

With regard to traffic, the PIDs are used on both uplinks and downlinks to identify connections of specified MAC QoS class from an ST to other STs in the same or different beams.

On the mesh uplinks, the PIDs are used in the IP/MPEG encapsulation (MPE) process. Each ST should be assigned a number of PIDs equal to the number of MAC QoS classes supported. In the encapsulation process the identity of the destination ST is captured via its MAC address in the header of the DSM-CC sections.

The usage of PIDs for traffic on mesh downlinks is similar to their usage in the access topology, i.e. for filtering of MPEG packets at the destination ST in order to re-assembly the original IP packets. PID filtering is followed by destination ST MAC address filtering (of DSM-CC sections) and then by IP filtering.

For the re-assembly process the PID should uniquely identify the source ST and the MAC QoS class. The source ST can be in principle be in any channel (from any beam) that can be established to the destination ST (as per mesh connectivity configured at a given time).

In the above context the mesh PIDs used for traffic may follow an addressing scheme of the format SQ, where S identifies the source ST and Q identifies the MAC QoS class. The scope of the PID is the channel, i.e. the same PIDs can be used by an ST in all authorised mesh channels (see bellow the mesh Channel_id description). The Q-field defines "virtual pipes" of given MAC QoS class within a channel. The size of such pipes should be configured based on the aggregation of Mesh Subscribers SLAs.

The number of PID values in the network is limited to 8192 values (13 bits), for both traffic and signalling, so the number of terminals that can be uniquely identified it rather small. Given the limited rates on DL TDMs, it is expected that in a practical mesh network only a limited number of terminals will establish connections within a channel ending in a given downlink TDM (the actual requirements for IP packet re-assembly is to have **unique PID values on any downlink TDM**).

The PID values should be configured in all relevant network components, in consistency with the PID plan defined for the entire INAP domain (by the NCC). Due to the multiplexing of traffic (from different STs) and signalling (from NCC) on the same downlink transport stream, the PID plan should be carefully designed.

DVB-RCS Identifiers

Group_id and **Logon_id**

Used as for the access topology**.**

Channel_id

The Channel_id is used in the mesh topology exactly as in the case of access topology, typically to identify resources associated with a destination, accessible from a given user beam. The destination is another user beam or a service beam (where the NCC is located). Channel_id is primarily used for the management of the uplink return link resources (by the Scheduler).

A ST can be configured with up to 16 Channel_ids. Channel_id "0" is typically reserved for the access channel (to the service beam, where ST-GW and NCC are located).

Population_id

Used as for the access topology.

In summary, in the case of mesh topology, with the exception of Population_id which is used only on the downlink, all other logical identifiers are used on both uplink and downlink.

D.2.2.1.3 Multiple PVC model

Multiple PVCs are required when multiple priorities are offered at IP layer.

Single PVC case

In the single PVC case the PVC is completely managed by the Satellite Bi-directional Channel (SBC) object that logically defines the full-duplex virtual link between the Satellite interface of the Terminal and the Satellite interface of the Gateway carrying IP packets (see TN1).

Figure D.4 illustrates how the IP packets are handled in the case of two IP layer priorities (high priority – HP and low priority – LP) and one PVC. The HP PDUs and LP PDUs are directed to the HP queue (HPQ) and LP queue (LPQ), respectively. The packets in the two queues are serviced by the ATM segmentation and scheduling function as follows:

Whenever there is an AAL5 PDU in the HPQ, it is always serviced before PDUs in the LPQ.

An AAL5 PDU (whether it is from HPQ or LPQ) must be entirely serviced by the ATM segmentation and scheduling function before another PDU can be serviced. Interleaved cells from different PDUs cannot be re-assembled, as illustrated in the bottom part of the figure.

These rules give IP service differentiation, while allowing the correct re-assembly of the ATM cells into an AAL5 PDU at the Gateway, but they also introduce extra delay and jitter, as an HP PDU cannot be transmitted before a scheduled LP PDU is fully serviced. This may not be compatible with real-time applications such as VoIP.

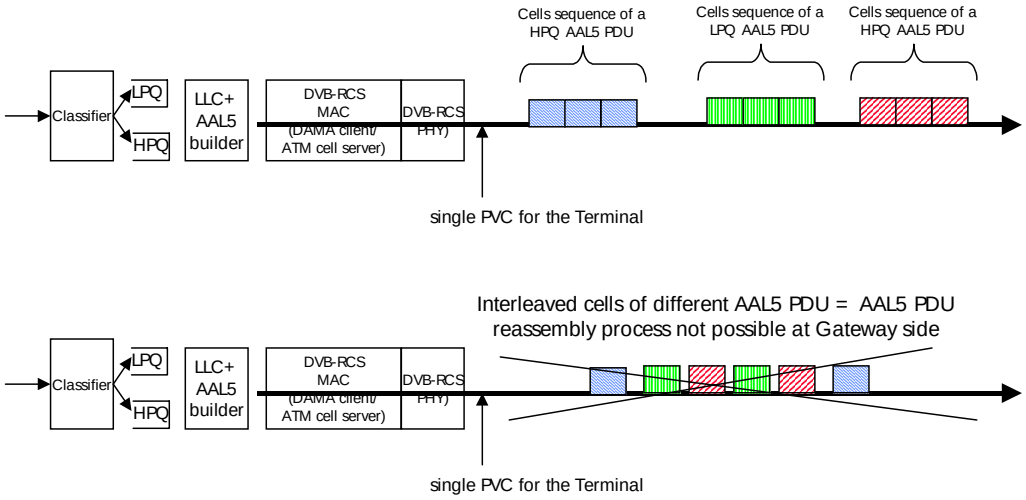


Figure D.4: Single PVC

Multiple PVC case

This multiple PVC case (two PVCs illustrated in figure D.5) is defined in order to cope with the limitations of single PVC case. The encapsulation and scheduling processes for the two IP priorities are independent of each other (they are PVC-specific) and so is the re-assembly at the Gateway side. This allows the interleaving of HP and LP PDUs (bottom part of the figure), with beneficial effect on delay/jitter for the HP IP packets.

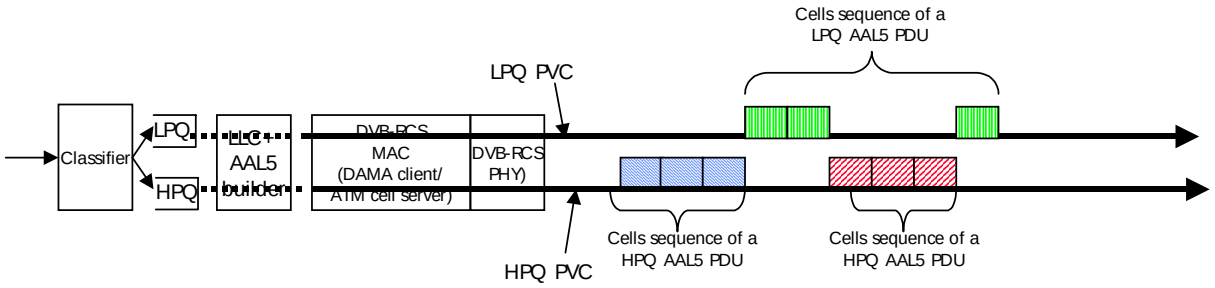


Figure D.5: Multiple PVC

The PVC used for LP traffic is typically configured at ST log-on (via a TIM message). The second PVC, for HP traffic, is configured right after the log-on through a SNMP "set" message.

D.2.2.2 Physical resources

Level 2 physical resources are organised in TDM carriers on the forward link and MF-TDMA carrier groups on the return link. Primarily defined for access topology, the uplink return link and the downlink forward link are also applicable to the mesh topology. The fact that different return link options have been suggested for access (ATM) and mesh (e.g. MPEG) has no or little importance with regard to level 2 mechanisms.

From level 2 viewpoint the forward link consists of continuous streams of MPEG2 TS packets, typically one TS per TDM. The Access Gateway should handle a number of TDMs, at least one for each user beam in the INAP.

By contrast, the return link has a frame structure associated with multiple MF-TDMA carriers, possibly of different rates. The frame structure is consistent with the DVB-RCS standard and consists of superframes, frames and timeslots. The standard provides a great flexibility in terms of frame and super-frame organisation, described by the superframe, frame and timeslot composition tables (SCT, FCT and TCT), collectively referred to as Burst Time Plan (BTP).

In a typical DVB-RCS system implementation:

All frames have the same duration and the same bandwidth (capacity).

A slot in a frame can be of one of the following types: overhead slot (i.e. CSC, ACQ, SYNC) or traffic slot (TRF). Frames with various combinations of overhead and traffic slots may be defined (via the BTP).

One slot carries one traffic burst (e.g. ATM, MPEG), or one CSC burst or 2 SYNC bursts.

The frame duration, while not defined by the standard, is typically set to 26,5 msec, which will provide a resource control granularity of 16 kbps (for ATM format).

The frame is the basis for the timing of all resource control processes (i.e. scheduling-related processes).

A BTP is used to define the slot composition of each UL RL transponder bandwidth. Terminals in a user beam will typically have access to one transponder bandwidth or to a portion of it.

In the case of access topology the entire transponder bandwidth (or capacity) is used to connect terminals to the access Gateway.

In the case of mesh topology, channels are defined within the transponder bandwidth, to allow terminals in a beam to access terminals in other beams or to access a Gateway/NCC. Mesh channels are defined in terms of slots on MF-TDMA carriers; their shape should match the connectivity configured in the OBP at any given time. There might be some restrictions in mesh channel definition, reflecting potential constraints in the OBP design/configuration. In addition to the physical resources, a mesh channel is also characterised by a source and a destination. The source is an uplink beam (or an UL RL transponder within a beam), while the destination is a downlink TDM. The channels accessible to an ST are uniquely identified by their Channel_ids.

Within the accessible channels a terminal can establish connections with other terminals (or Gateway/NCC), located in other beams and receiving the downlink TDMs defining channels' destinations.

D.2.3 Layer 2 Functionality

Level 2 functionality will be defined in the user plane and control plane, for both access and mesh topologies.

D.2.3.1 User plane functionality

Level 2 functionality in the user plane includes transport functions that need to be implemented in all satellite network components, i.e. ST, GW (both Gateways and ST-GW), ST-NCC, satellite. The satellite also needs to implement switching function. The transport/switching functions include:

- Exchange of packets with the network layer:

  Two types of network layer packets are being considered for the study, namely IP packets and Ethernet packets (frames). However, the study only concentrates on QoS for traffic in IP format.

- Packet formatting / exchange with the physical layer:

  The formats on the physical layer are consistent with the modulation/coding schemes defined in the DVB-RCS standard.

- Encapsulation of the network layer packets in native format (IP, Ethernet) into the formats supported on the MAC layer of the air interface (on FL, RL):

  On the **forward link** the MAC format is MPEG.

The encapsulation of the IP and Ethernet packets into MPEG packets takes place in the IP/DVB Gateway in the Gateway, which is configured with IP/MAC/PID triplets (in the case of IP packets) and MAC/PID twoplets (in the case of Ethernet packets). It is based on the Multi-Protocol Encapsulation (MPE): the packets are encapsulated in DSM-CC sections with the MAC address of the destination (ST in the case of IP packets, host in the case of Ethernet packets), and the sections are segmented in MPEG packets with the configured PID value. The source MAC address in the Ethernet frame header is that of the last IP component on the forward path within the Gateway.

NOTE: Please note that in the case of management traffic (e.g. OA&M) the destination address is that of the ST even in the case of Ethernet packets.

On the **return link** the MAC format is ATM, and the encapsulation of IP packets is based on AAL5.

The Ethernet packets may also be encapsulated via AAL5, after stripping off the Ethernet frame header. Alternatively, if Ethernet packets are used in conjunction with PPPoE access mode, the multi-protocol encapsulation over AAL5 can be used (as per RFC 2684, see Bibliography) for the encapsulation of both native IP packets and PPPoE packets.

On the **mesh links** (both uplink and downlink) the format is MPEG.

The encapsulation process is the same as for the forward link, but it takes place in terminal. The MAC address in the DSM-CC sections is that of the destination host (user traffic) or destination ST (OA&M traffic).

- Multiplexing of traffic with FLS signalling (generated in layer 2 format);

  The FLS signalling is part of the control plane (see below).

  In the case of access topology the multiplexing takes place in the IP/DVB Gateway, which includes a DVB MUX function. In the case of mesh topology the multiplexing takes place in the OBP; it is a multiplexing of selected MPEG packets from different uplink transport streams into downlink transport streams.

- Switching of layer 2 packets in the OBP;

  The switching is of fast circuit type and is done according to the frequency and time position of the incoming MPEG packets. Packet identifier (PID) and other identifiers used by the access layer (see clause D.2.2.1) have no role in on-board switching. The circuits configured in the switch should reflect the connectivity and the associated physical resources valid at a given time (as defined via channels). Since the switching takes place at layer 2, an interface between layer 2 and the physical layer needs to be implemented in the OBP.

- Re-assembly of native network layer packets in their original format;

  The re-assembly follows the same protocol stack as that used in the encapsulation process.

  In the case of access FL and mesh (MPEG format) the re-assembly takes place in ST, based on MPE. The packets are filtered per PID, then per destination MAC address (to re-assembly the DSM-CC sections) and then per IP address (to recover the IP packets) or per source MAC address (to recover the Ethernet frames).

  In the case of access RL (ATM format) the re-assembly takes place in Gateway, based on AAL5.

  From the above description we can say that there is no QoS awareness at the access layer in the user plane.

D.2.3.2 Control plane functionality

Level 2 functionality in the control plane includes functions associated with terminal access (logon) to satellite network, link control, resource control, table generation and distribution. Such functions need to be implemented (to various degrees) in all satellite network components, i.e. ST, GW (both Gateway and ST-GW), NCC-ST, satellite.

Logon process

The logon process is the process by which a ST gains access to the return link of the interactive network. It relies on both RL (CSC, SYNC) and FL (TIM, CMT) signalling. The following events take place during a successful logon process:

- ST validation, authentication, registration with the network.
- Host authentication (with a RADIUS server).

  For RADIUS messages the insecure IP OA&M path is used (after having the OA&M PVC assigned to the ST).

- Connection admission / resource reservation.
- RL synchronisation acquisition.

  The coarse synchronisation is acquired by using CSC / ACQ burst(s), while the fine synchronisation relies on SYNC bursts. The frequency/time errors are included in CMT replies.

- Triggering of collection of accounting information.

  In addition to the DVB-RCS signalling mentioned above (CSC, SYNC, CMT), the logon process also relies on TIM messages (from Gateway/NCC to ST), carrying logical identifiers (Group_id, Logon_id, Channels_ids, ST IP addresses, PID values, PVCs) and accept/reject messages.

Link control

It includes:

- Fine synchronisation maintenance.

  Relies on periodic SYNC bursts and time/frequency errors in the CMT replies. The time/frequency errors are measured either on ground (access topology) or on-board (mesh topology).

- Uplink power control (in ST).

  Relies on periodic SYNC bursts and CMT replies including power or En/No measurements.

- Uplink power control (in Gateway).

  Optional. May use Radiometers.

Signalling generation/distribution and interface with the physical layer

Layer 2 signalling includes the RL signalling and FL signalling.

The formats for RL signalling (CSC, ACQ, SYNC) transmission on the physical layer are consistent with the modulation/coding schemes defined in the DVB-RCS standard.

The FL signalling includes the DVB tables, both standard tables and RCS-specific tables (for the definition of these tables please refer to the DVB-RCS standard). They are carried in MPEG TS packets, multiplexed with the traffic packets. From level 2 viewpoint the DVB tables originate in the IP/DVB Gateway. Table content is either provided by the NCC or generated within the RLSS. However, in the case of mesh topology, the content of the CMT originates in the OBP, where the power/time/frequency measurements are performed. The results of measurements are encapsulated in specific MPEG packets, which are switched to the appropriate DL TDMs, according to the switching configuration in OBP. The resources required for signalling should be accounting for in the definition of the switching tables. In the Gateway/NCC these specific MPEG packets are PID-filtered in the same way as the traffic packets. The Gateway/NCC then builds the CMT messages for transmission to STs.

The transmission of DVB tables requires FL resources. The most resource-intensive is the TBTP, which is transmitted every frame (like the CMT). Other tables are transmitted with a lower periodicity or as needed (e.g. TIM table). The consequence is a variable load from the DVB tables. As the FL rate is fixed, the capacity available for user traffic is also variable. This has implications on implementation and configuration of level 2 (DiffServ) mechanisms on the FL (in the IP/DV Gateway), as discussed in clause D.1.

The formats for FL signalling transmission on the physical layer are as for the user traffic.

The extraction of the DVB tables at the terminals is based on PID filtering.

Resource control

Layer 2 resource control refers to the management (assignment) of the MAC physical resources for the connections that have been admitted and set-up, based on predefined rules. This is a function associated with schedulers. The schedulers are only concerned with the management of the uplink resources; the on-board processor ensures that the resources assigned on the uplinks will also be available on the corresponding downlinks (as per the connectivity tables). In order for this to happen the resources need to be organised (in schedulers) to reflect the connectivity tables.

On the forward link

On the forward link the resources are directly provisioned based on a resource sharing TDM scheme.

The management of the forward resources is performed by the ACSS within the Gateway and is based on "virtual" capacity pipes (typically associated with service providers) established along defined forward paths (as per system connectivity). The packets are fed into these pipes as they become available from the network (IP) layer, with no terminal-based or service–based scheduling discipline. The order of packets established at the IP layer is not changed after MPEG encapsulation. The encapsulation takes place in the IP/DVB Gateway, which also multiplexes the user packets with signalling packets (DVB standard tables and DVB-RCS tables). The process is illustrated in figure D.2 for a typical IP/DVB Gateway (e.g. SkyStream SMR 25/26).

NOTE: Some of the functions of the ACSS are currently implemented in the Connection Manager (CM) subsystem, also referred to as Traffic Manager (TM). ACSS is an upgraded CM, to support dynamic SLA management.

On the return link

On the return link the resources are subject to contention between STs. The contention is resolved by the MAC Scheduler within the RLSS. This is part of the RL dynamic resource control covered in a separate clause (clause D.2.4) given its importance in a DVB-RCS system and its relevance to QoS provisioning.

D.2.4 Return (Uplink) Link Dynamic Resource Control

D.2.4.1 Overview

Dynamic resource control consists in assignment of resources (slots) to terminals based on their requests and limit values negotiated/set during connection establishment. The assignments are conditioned by the availability of resources (capacity) within defined return channels (as per system connectivity). The assignment is the responsibility of the MAC Scheduler, which implements a Demand-Assignment Multiple Access (DAMA) protocol.

The above process applies to return links in access topology, but also to the uplink in mesh topology, therefore the process is alternatively referred to as uplink scheduling.

The MAC Scheduler is located on ground for the access topology (in the Access Gateway). For the mesh topology it can be located on ground or on-board the satellite.

The uplink scheduling consists of processes taking place in the Scheduler and in terminals, namely:

- Calculation of capacity requests in terminals.
- Transmission of capacity requests to the Scheduler (request signalling).
- Calculation of capacity assignment to terminals.
- Allocation (placement) of the assigned capacity.
- Generation and transmission of TBTPs carrying the allocations to terminals.
- Distribution (within terminals) of the allocated capacity to the end users and their applications (consistent with ST MAC queue architecture and service discipline).

Each of these aspects will be addressed in the next sections and the available options will be analysed.

Uplink scheduling for access and mesh topologies is similar but not identical. The basic processes are the same, but there are differences resulting from the resource organisation in the Scheduler for the two topologies. The most important difference is related to the number of channels that need to be supported: one for access, several for mesh.

The role of the Scheduler in QoS provisioning will be analysed in the context of the overall approach taken to IP-QoS. In this section it suffices to say that IP-QoS is based on the DiffServ architecture and mechanisms and that it is primarily dealt with at network layer. However, the classes of service are propagated into the MAC layer (MAC QoS classes) and the Scheduler is capable of providing differentiated access to the uplink resources by proper mapping of MAC QoS classes to capacity types.

D.2.4.2 Resource organisation in the Scheduler

For the purpose of switching, the uplink resources in a beam (i.e. slots on MF-TDMA carriers) are organised in physical channels, that are switched on-board the satellite to different destinations (downlinks in user beams for mesh or a service beam for access). All channels have the same source – the uplink beam under consideration. For the access reference model there is only one channel in an uplink beam.

The channels reflect the connectivity configured at a given time in the on-board switch. The assumption is made that the on-board switch allows a channel to comprise any number of slots distributed anywhere in the uplink frame. Figure D.6 illustrates an uplink with four channels, comprising compact blocks of slots laid over the slots/carriers defining the MF-TDMA frame (see area discussion below). Two carrier rates have been illustrated with a 2:1 ratio and the assumption was made that the MCD on board the satellite can change rate in the mid carrier.

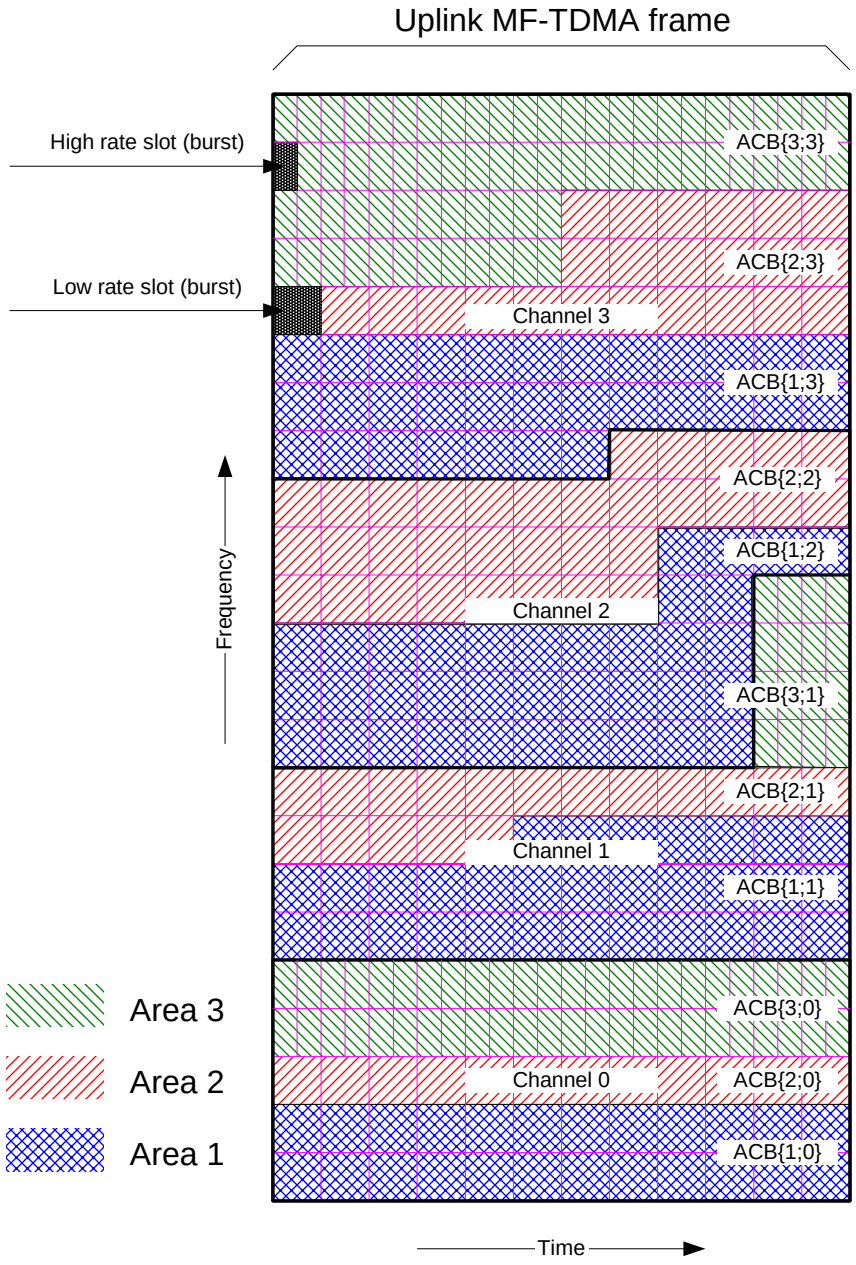


Figure D.6: Areas & Channels laid over a MF-TDMA frame

For the purpose of scheduling the total capacity in an uplink beam (belonging to the IN) is divided in areas, with each area occupying a part of one or more channels. An area corresponds to a logical grouping of capacity serving a particular group of terminals and having the following properties:

- All carriers / part carriers in an area must have the same symbol rate and slot timing.
- A given terminal belongs to one, and only one, area.
- Each area is owned, and scheduled, by a single INAP. An INAP can have multiple areas in a given beam (e.g. corresponding to different rates, or each served by a different forward link), and a beam can contain areas belonging to several INAPs.

- An area comprises a number of blocks of slots, where each block corresponds to the part of that area in a given channel. Each such block is conveniently called an Area – Channel Block (ACB)[5].

- Each area is scheduled independently of any other area, and a single scheduler must schedule all parts of the area. The same scheduler can schedule several areas, if they are under the same administrative domain (INAP).

Each channel does not necessarily support every area and conversely each area does not necessarily support every channel (this depends on the needs of the access provider for each area). We can consider areas and channels as subdivisions of the uplink beam capacity of equal hierarchical level, with each ACB representing a conjunction of an Area / Channel pair. Of the three areas illustrated in figure D.6, areas 1 and 3 have capacity in all channels, while area 2 is only present in channels 0, 1 and 3.

Channel_id (as used by the terminals in the beam and by the Scheduler) will match the channel number in figure D.6. This assumes that Channel_id is beam-specific and not ST-specific, i.e. a given Channel_id will designate the same physical channel for all STs in a given area in a beam. For the areas illustrated in figure D.6 this means that Channel_2 is not implemented in the STs belonging to area 3.

The ACB (i.e. the intersection of an area and a channel) is the basic building block for the schedulable frame structure. Specific ACBs are referenced as ACB{Area; Channel}, where {Area; Channel} are the ACB co-ordinates (unique for each ACB). As ACBs are divisions of channels for which no shape constraints have been considered, one can have ACBs with shapes creating scheduling problems (more accurately slot placement problems). ACB{3;1}, for example, has 6 high rate slots but an ST can only use up to 3 slots due to collision resulting form the particular shape of the ACB. This problem can be accounted for in at least two ways:

- At scheduling level, by imposing constraints on the size/shape of ACBs. One can for example lower limit an ACB to a carrier-worth of contiguous slots on the same or consecutive carriers. This creates the premises for an ST to transmit at its maximum rate (i.e. area rate).

- At connection control level, by configuring in the ACSS (or, in general, in the network component responsible for call admission) the maximum rate an ST can use in a given channel (as given by the number of non-overlapping time slots in the corresponding ACB). Even if the size/shape of an ACB is accounted for at connection level, Scheduler performance degradation are expected (due to collisions), unless some constraints are imposed on the size / structure (slot position) of ACBs.

D.2.4.3 Scheduling hierarchy

Conceptually, MAC scheduling can be considered as a hierarchical process, that reflects the organisations of RL resources.

D.2.4.3.1 Access topology

In access topology there is only one channel per beam. The subscribers associated with terminals in the beam can belong to different SPs, as per the access reference architecture. An SP is associated with a segment. Each segment within each area "owns" a subset of area capacity and has guaranteed access to that capacity. Depending on the system configuration, the segment may also have access to area capacity that is not owned by any segment (area pool), on an "as available" basis.

A given MAC Scheduler can support a number of allocation areas, each of which supports a number of assignment segments, each of which supports a number of STs, as shown in the scheduling tree in figure D.7. An ST can only belong to one segment in one area.

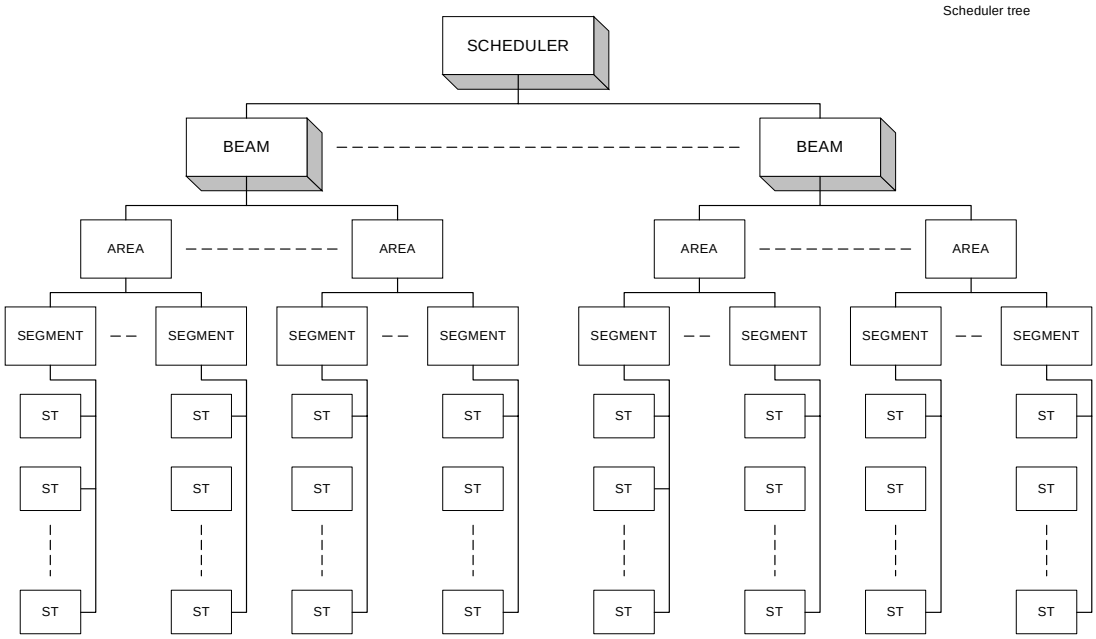


Figure D.7: MAC scheduling hierarchy – access topology

D.2.4.3.2 Mesh topology

The subscribers associated with the mesh terminals are of the same service provider, associated with the mesh network (they are part of a VPN). In this context there is no need to support segments. The scheduling hierarchy in this multi-channel single-segment environment is shown in figure D.8.

A given MAC Scheduler can support a number of allocation areas, each of which having capacity in a number of channels. A ST only belongs to one area, but may request capacity / receive assignments in different channels.

The architecture in figure D.8 and the RL resource organisation does not preclude the support for segments in the case of mesh topology (if needed for other reasons, such as grouping of terminals for billing and/or accounting purposes). A segment can have capacity in several ACBs, but there is no association between segment capacity in an ACB (which is a scalar) and physical resources.

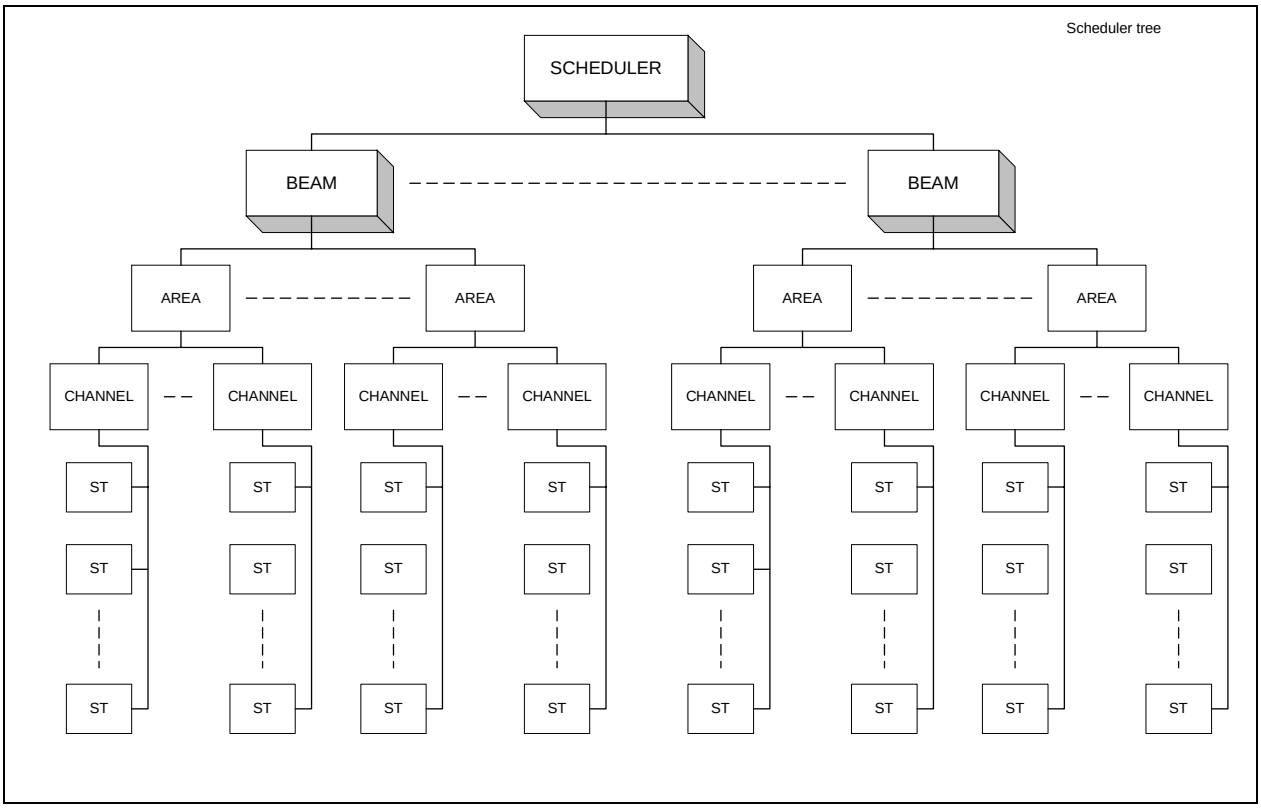


Figure D.8: MAC scheduling hierarchy – mesh topology

D.2.4.4 DVB-RCS capacity types

The types (categories) of capacity that can be assigned to an ST are consistent with the DVB-RCS standard. They are:

- CRA, or Static Rate (SR).
- RBDC, upper-bounded by MaxRBDC.
- VBDC.
- FCA.

The standard also defines an Absolute VBDC (AVBDC), which can be used in the MAC Scheduler to recover from requests/allocation losses.

Capacity types are vital to QoS support at MAC layer, therefore they are describes in detail below, in order of their priority (highest to lowest). Any given terminal can be assigned one or a mix of the four capacity types.

D.2.4.4.1 Constant Rate Assignment (CRA)

CRA is static rate capacity. It is **static** capacity, as it is not subject to dynamic requests; rather, it is rather set at logon time and presumably remains constant for the duration of the session, unless it is renegotiated during the session (for example using COPS). It is **rate** capacity, defined in slots/frame, as it is granted in full every frame, provided that the total CRA per ST is within the maximum ST transmit capability and does not exceed the area rate, and that the sum of all CRAs for all STs in a segment/area is within the segment/area capacity.

CRA capacity is set equal to the value in Subscriber SLA, typically negotiated off-line. Within the context of the overall QoS management architecture proposed in this study, it can also be signalled during session establishment (e.g. via SNMP or COPS). CRA needs to be configured in the GMSS (in ST data base) and the RLSS (in ST table within the MAC Scheduler), either statically or dynamically.

CRA capacity is said to be **reserved**, in the sense that a terminal does not need to request it every frame. It is always assigned to a terminal, whether the terminal has traffic to send or not.

CRA is typically used for the highest priority delay-sensitive user traffic that requires a fixed guaranteed rate. It may also be used for signalling.

D.2.4.4.2 Rate Based Dynamic Capacity (RBDC)

This capacity category is used for high priority variable rate traffic that can tolerate the MAC Scheduler dynamic response time latency. RBDC is dynamic rate capacity (in slots/frame) granted in response to dynamic requests from a ST to track the instantaneous traffic rate. A maximum rate limit (MaxRBDC) is set within the limits of the Booked Rate set by the SLA and is configured in the MAC Scheduler and the GMSS. As with CRA, MaxRBDC can be negotiated off-line and stored in GMSS data base or can be signalled during session establishment.

RBDC capacity is said to be **guaranteed** if the sum of CRA and MaxRBDC values configured in the MAC Scheduler for each ST are constrained to the terminal physical transmission limit and to area rate, **and** if the total (CRA + MaxRBDC) values for all STs in the segment/area is within the segment/area capacity. If these conditions are met, the RBDC is appropriate for traffic that requires bandwidth guarantees.

In case of overbooking the latter condition above is not met and therefore no absolute guarantees can be provided that the requests will be granted in the expected frame. Various strategies can be used concerning the requests in excess of MaxRBDC. They can be ignored or they can be granted as VBDC (if available), but not necessarily in the expected frame. In the latter case the packets queued in the ST might eventually be transmitted, but with additional delay, so jitter will be induced. Overbooking may be used by some network operators in order to increase capacity utilisation. It relies on the fact that, due to traffic variability, not all STs using RBDC transmit packets at maximum rate (i.e. MaxRBDC). In this case the capacity is guaranteed on a statistical base.

The actual amount of RBDC granted in any superframe is controlled by dynamic requests from the ST to the MAC Scheduler, each request being for the full RBDC rate currently needed (while accounting for the requests in progress). Each request overrides all previous requests from the same terminal, and is subject to a configurable time-out period, to prevent a terminal anomaly resulting in a hanging capacity assignment.

D.2.4.4.3 Volume Based Dynamic Capacity (VBDC)

This capacity category is used for traffic that can tolerate delay jitter, such as the Best Effort (BE) class of the Internet traffic. VBDC capacity is provided in response to dynamic volume requests from the ST to the MAC Scheduler (a volume request is for a given number of slots with no time constraint), which are accumulated at the MAC Scheduler in a queued volume count, Q, per terminal. Algorithmically, if each request is for R cells worth of capacity, then it is added directly to Q, i.e. Q => Q+R. In any given superframe, capacity is assigned up to the current value of Q, and the amount assigned, A, subtracted from Q, i.e. Q => Q - A.

In general VBDC capacity is not guaranteed. It is assigned as **best effort** capacity within the available resources, after satisfying the total CRA and RBDC capacity components. The amount of VBDC capacity assigned to a terminal can be limited to a MaxVBDC value, based on rules set by the network administrator or on results from measurements on traffic.

A **guaranteed VBDC** (G-VBDC) or High Priority VBDC (HP-VBDC) capacity can also be defined, by setting a minimum value for VBDC (MinVBDC) per ST and configuring it in the MAC Scheduler and ST. VBDC capacity up to MinVBDC or Q value (whichever is less) will be granted every superframe, and so VBDC capacity up to MinVBDC is treated as a third form of guaranteed capacity along with CRA and RBDC. Note, however, that it does not obey the RBDC rules: there is no special request for this capacity (it is part of the VBDC request) and there is no time-out feature, so it cannot be reset. VBDC requests for capacity in excess of MinVBDC will compete for the balance of (standard) VBDC capacity with other STs. This makes it very attractive for implementing the Internet AF traffic classes.

D.2.4.4.4 Free Capacity Assignment (FCA)

This capacity category is assigned to STs on an "as available" basis, from capacity that would be otherwise unused. This is automatic and does not involve any capacity request from the terminal to MAC Scheduler and the ST has no say in FCA assignment. In the basic MAC Scheduler design, each terminal is limited to one FCA slot per superframe, but alternatively more than one slot can be freely assigned, where the ratio of terminals to free capacity makes it meaningful. A terminal may qualify or not for FCA, based on a configuration parameter in the terminal profile, which is made available to the MAC Scheduler.

FCA can improve the MAC Scheduler performance (throughput), especially in low loading conditions, but it can introduce jitter, therefore its use should be limited to applications that are jitter-tolerant. FCA is normally used to supplement VBDC.

D.2.4.4.5 Mapping of MAC QoS classes into capacity types

In clause D.1.2 it was suggested that at least one MAC QoS class should be used for each DiffServ class of service type (i.e. EF, AF and BE). If we consider the network management (NM) traffic in a MAC class of its own, there will be 4 MAC QoS classes of service, namely NM, EF, AF and BE (see figure D.3).

In principle any MAC class of service can be mapped in any type or combination of types of capacity, by properly setting the queue attributes (which define the values for authorised types of capacity for a given queue).

From the above four capacity types, CRA and RBDC and G-VBDC can offer guaranteed capacity, while VBDC and FCA are of best effort type; the total guaranteed dynamic capacity is referred to as **booked capacity** or **booked rate**. From another prospective, RBDC and VBDC are based on capacity requests, while CRA and FCA involve no request and therefore may reduce the delay (at the expense of additional jitter in the case of FCA).

In the case of dynamic types of capacity (i.e. RBDC and VBDC) the delay performance can be expressed via the scheduling latency, defined as the difference between the time a capacity request has been made by a ST and the time the assignment can be used by the ST to send traffic. The Minimum Scheduling Latency (MSL) is given by:

$$MSL = Tpd(FL) + Tpd\ (RL) + processing\ time$$

where Tpd(FL) + Tpd(RL) are the propagation times on the forward path and return path (typically 280 ms each for a GEO satellite). The Gateway processing (extraction of capacity requests and generation of assignments to STs) takes typically 4 frames, while ST processing (mainly extraction and dispatching of the assignments) takes 2 frames. A frame duration of 26,5 ms gives an MSL of 666 ms.

Undoubtedly CRA offer the best performance in terms of both delay and delay variation, as it is not only guaranteed but also reserved. The delay is given by the Tpd(FL) + processing time, or about 386 ms. The average jitter is about a frame duration. CRA is best suited for jitter and delay sensitive applications such as voice applications (e.g. VoIP), but it is an expensive type of capacity, not always affordable.

In the case of RBDC the delay is given by Tpd(FL) + MSL, or 946 ms, while the jitter is the same as for CRA, i.e. one frame duration.

For both CRA and RBDC, the delay and jitter are independent of load, as the two types of capacity are guaranteed.

In the case of VBDC the delay is given by Tpd(FL) + MSL + queuing in ST. The queuing time depends on the load; the delay is therefore lower limited by Tpd(FL) + MSL, or 946 ms, but increases with system loading, reaching seconds at 90-95 % load. The jitter too varies function of the system loading and can be of several frames. The use of FCA may improve the delay (even below 946 ms), but adds extra jitter.

RBDC with absolute capacity guarantee can successfully be used for applications sensitive to packet losses and/or jitter but with no critical delay requirements, provided that the application generates a sustained stream of packets, thus providing regular request opportunities. FCA should be avoided in such case; it is true that it can speed up the traffic transmission but it also adds jitter. Please note that FCA has no impact on applications using CRA.

If RBDC is overbooked, only statistical guarantee can be provided. Depending on the overbooking factor and traffic characteristics, the throughput may not be affected, but delay variation will be induced.

VBDC, combined with FCA, can be used for applications tolerant to packet loss, delay and jitter. Where a minimum rate is desired (in order to reduce the latency) a guaranteed VBDC component can be added.

In general higher priority classes of service are associated with guaranteed capacity (CRA, RBDC, G-VBDC), while lower priority classes are predominantly given best effort capacity (VBDC, FCA). As already mentioned, G-VBDC is a guaranteed form of VBDC (up to a MinVBDC limit), which is very suitable to AF class forwarding needs. A potential mapping of the MAC QoS classes (MAC queues) into capacity types is suggested in Table D.1. The table also shows the correspondence of these classes with the DiffServ classes of service / precedence and makes reference to the assumed traffic conditioning at IP level.

Table D.1: Example of classes of service mapping into capacity types

DiffServ CoS / Precedence	MAC Queue	Type of capacity
LNM / INM	NM	CRA or G-VBDC + VBDC, with proper traffic conditioning at IP layer (shaping, no discarding).
EF	EF	CRA or RBDC or CRA + RBDC, function of service requirements.
AF1-4	AF	G-VBDC + VBDC + FCA (shaping with short buffer size, discarding).
BE	BE	VBDC + FCA.

The use of absolute guaranteed capacity should be considered not only from the point of view of service requirements, but also from the point of view of ST and MF-TDMA constraints (i.e. rates) and ST subscriber model. On a low rate carrier the number of slots is rather limited (e.g. 24 slots at 384 kbps user rate and 26,5 ms frame duration). This may result in a number of problems, such as:

- waste of capacity (when all traffic from a ST requires guaranteed service and the traffic exhibits some variability);
- too much fragmentation of capacity:
  - in the case of multiple channels per ST (e.g. mesh);
  - in the case of service agreements with multiple SPs per ST (if allowed by the Subscriber SLA model).

D.2.4.5 Capacity requesting mechanisms

There are two fundamental mechanisms for capacity requests sent by the terminals to the Scheduler:

- **In Band Requests** (IBR), in which the requests are signalled in a supplemental header of all data cells sent by the terminal. This header is stripped off by the receiving demodulator and forwarded to the scheduler.
- **Out of Band Requests** (OBR), in which the requests are signalled independent of the traffic slot assignments.

DVB-RCS standard makes provision for both IBR and OBR. The IBR is associated with the prefix method and with the Data Unit Labelling Method (DULM), while the OBR is associated with the minislot method (relying on SYNC bursts). In the case of prefix method and minislot methods the requests are encapsulated in a special Satellite Access Control (SAC) field, together with other RL messages (M&C messages, Logon_id, Group_id). In the case of DULM the requests are sent in regular traffic bursts.

For the access topology the prefix method will be used (IBR), combined with the minislot method (OBR) with the SYNC in pre-assigned mode. The SYNC slots are primarily assigned to an ST in order to satisfy RL UPC and synchronisation needs, typically one SYNC every 32 frames (0,884 sec at 26,5 ms frame duration). The role of OBR is therefore merely to speed up the initial access to return link capacity.

For the mesh topology, the minislot method (SAC field in SYNC burst) is used. The number of request opportunities can be increased by exploiting the variable SYNC burst structure and the SAC field structure (section 3.4.3.2.3 in the DVB-RCS standard). The number of requests per SYNC burst can vary from 2 to 4, function of the coding scheme/rate (that determines the useful payload size) and of the optional sub-fields in the SAC field that are used (system dependent). A terminal can thus make up to four independent requests at a time, either for different types of capacity or for different channels. When the scheduling latency is a concern, the periodicity of SYNC slots assigned to a ST may need to be increased, and this will negatively impact the signalling overhead in the MF-TDMA frame.

The decision on how the assigned SYNC bursts are used should reside with the terminal, based on rules set by Gateway/NCC and dependent on terminal traffic loading conditions (e.g. MAC queues condition). Such rules could be based on fairness principles or could define priorities per request type, per destination and/or other criteria (e.g. queue status or other monitored parameters). A request for RBDC should take precedence over a request for VBDC. Some destinations may have priority with respect to other destinations. A round robin mechanism (weighted or not) could be considered for example, for sharing the request opportunities among different logical channels.

D.2.4.6 Outline of the capacity scheduling process

The scheduling process is similar for the access and mesh topologies. However, they are not identical, as a result of different scheduling architectures. The scheduling process will be reviewed first for the access topology, and then the differences for the mesh topology will be highlighted.

D.2.4.6.1 Access topology

The overall capacity assignment process comprises a number of separate but dependent processes, some performed by the terminal and others by the Scheduler. The following is an outline description of this overall process in terms of the individual processes in the order of execution.

- Traffic arriving at the terminal input from the STs is queued into a number of queues. The first set of queues is at the IP layer and the second at the MAC layer. Their number depends primarily on the classes of service / types of traffic supported and the approach to QoS implementation (system-specific). From the point of view of Scheduler operation only the MAC layer queues are important.

- Based on the current MAC queues status and on queue attributes, the terminal issues RBDC and/or VBDC requests after allowing for requests already issued and for constraints imposed by the congestion control system (if any). The requests are issued when needed and are frame synchronized.

- On arrival at the Scheduler, the capacity requests, tagged with the identity of the source terminal, are buffered or stored in appropriate buffers/queues. The source terminal is derived from the burst (carrying the request) assignment in the superframe.

- At the start of each superframe the Scheduler uses the buffered requests received in the last scheduling interval, plus unsatisfied requests from prior scheduling intervals, to assign capacity to terminals.

- The Scheduler builds TBTP update messages corresponding to the assignments from the previous process, and transmits them on the terminal's forward link every superframe.

- On receiving the TBTP update messages, the terminal decodes them and dispatches traffic cells from the MAC priority queues for transmission in the allocated time slots of the specified UL superframe number. In parallel, it updates the queue status to reflect the committed traffic.

D.2.4.6.2 Mesh topology

The single, most important difference (with respect to access topology) in the scheduling process for mesh topology results from the scheduling multi-channel architecture.

A terminal makes separate requests for capacity in each channel in the area (the Channel_id is coded in the SAC field), but the processing of requests in the Scheduler cannot be made independently per channel, due to terminal / MF-TDMA frame constraints (an ST can only transmit in one slot at a time and the maximum number of slots is limited by the frame rate). The assignment process described for the access topology needs to be complemented by an allocation (slot placement) process. The allocation process is responsible for preventing traffic slots assignments in different channels to collide (with either traffic slots or overhead slots, e.g. SYNC).

Slot placement might also be required in order to reduce the jitter (in the case of jitter sensitive traffic), if the superframe duration is relatively long. The current superframe duration of 26,5 ms is considered sufficiently small, so that the jitter induced is acceptable for all applications/services considered in this study.

There are three basic approaches (models) to slot allocation:

- approach 1: the allocation is a different process down the scheduling stream (after assignment and independent of assignment);

- approach 2: the allocation and assignments are done in the same time;

- approach 3: the guaranteed assignments are allocated at the time of assignment, the non-guaranteed capacity is allocated after the assignment process.

From the above, approach 3 is the most promising, as it allows the slots collided during the placement of guaranteed assignments to be used for non-guaranteed assignments.

The need for placement algorithms increases substantially the scheduler processing requirements.

Annex E (informative): Bibliography

- ETSI TS 102 295: "Satellite Earth Stations and Systems (SES); Broadband Satellite Multimedia (BSM) services and architectures; BSM Traffic Classes".
- ETSI TS 101 192: "Digital Video Broadcasting (DVB); DVB specification for data broadcasting".
- ETSI TS 185 001: "Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); Next Generation Network (NGN); Quality of Service (QoS) Framework and Requirements".
- ETSI TR 101 984: "Satellite Earth Stations and Systems (SES); Broadband Satellite Multimedia; Services and Architectures".
- ETSI TR 101 985: "Satellite Earth Stations and Systems (SES); Broadband Satellite Multimedia; IP over Satellite".
- ETSI TR 102 157: "Satellite Earth Stations and Systems (SES); Broadband Satellite Multimedia; IP Interworking over satellite; Performance, Availability and Quality of Service".
- ETSI TS 123 107: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); Quality of Service (QoS) concept and architecture (3GPP TS 23.107 version 6.4.0 Release 6)".
- ITU-T Recommendation E.360: "QoS routing and related traffic engineering methods".
- ITU-T Recommendation E.361: "QoS routing support for interworking of QoS service classes across routing technologies".
- ITU-T Recommendation E.860: "Framework of a service level agreement".
- ITU-T Recommendation I.112: "Vocabulary of terms for ISDNs".
- ITU-T Recommendation G.1010: "End-user multimedia QoS categories".
- ITU-T Recommendation Y.1291: "An architectural framework for support of Quality of Service in packet networks".
- ITU-T Recommendation Y.1221: "Traffic control and congestion control in IP based networks".
- ITU-T Recommendation Y.1541: "Network performance objectives for IP-Based services".
- IETF RFC 791: "Internet Protocol; Darpa Internet Program; Protocol specification".

NOTE: Available at: http://www.ietf.org/rfc/rfc791.txt

- IETF RFC 1349: "Type of Service in the Internet Protocol Suite".

NOTE: Available at: http://www.ietf.org/rfc/rfc1349.txt

- IETF RFC 2205: "Resource ReSerVation Protocol"

NOTE: Available at: http://www.ietf.org/rfc/rfc2205.txt

- IETF RFC 2309: " Recommendations on Queue Management and Congestion Avoidance in the Internet".

NOTE: Available at: http://www.ietf.org/rfc/rfc2309.txt

- IETF RFC 2386: "A Framework for QoS-based Routing in the Internet"

NOTE: Available at: http://www.ietf.org/rfc/rfc2386.txt

- IETF RFC 2676: " QoS Routing Mechanisms and OSPF Extensions"

NOTE: Available at: http://www.ietf.org/rfc/rfc2676.txt

- IETF RFC 2684: " Multiprotocol Encapsulation over ATM Adaptation Layer 5"

NOTE: Available at: http://www.ietf.org/rfc/rfc2684.txt

- IETF RFC 2748: "The COPS (Common Open Policy Service) Protocol".

NOTE: Available at http://www.ietf.org/rfc/rfc2748.txt

- IETF RFC 2749: "COPS Usage for RSVP" .

NOTE: Available at: http://www.ietf.org/rfc/rfc2749.txt

- IETF RFC 2750: "RSVP Extensions for Policy Control".

NOTE: Available at: http://www.ietf.org/rfc/rfc2750.txt

- IETF RFC 2751: "Signalled Preemption Priority Policy Element".

NOTE: Available at: http://www.ietf.org/rfc/rfc2751.txt

- IETF RFC 2753: "A Framework for Policy-Based Admission Control".

NOTE: Available at http://www.ietf.org/rfc/rfc2753.txt

- IETF RFC 2990: "Next Steps for the IP QoS Architecture".

NOTE: Available at http://www.ietf.org/rfc/rfc2990.txt

- IETF RFC 2996: "Format of the RSVP DCLASS Object".

NOTE: Available at http://www.ietf.org/rfc/rfc2996.txt

- IETF RFC 3052: "Service Management Architectures Issues and Review".

NOTE: Available at: http://www.ietf.org/rfc/rfc3052.txt

- IETF RFC 3084: "COPS Usage for Policy Provisioning or COPS-PR".

NOTE: Available at: http://www.ietf.org/rfc/rfc3084.txt

- IETF RFC 3168: " The Addition of Explicit Congestion Notification (ECN) to IP".

NOTE: Available at: http://www.ietf.org/rfc/rfc3168.txt

- IETF RFC 3260: "New Terminology and Clarifications for Diffserv".

NOTE: Available at: http://www.ietf.org/rfc/rfc3260.txt

- IETF RFC 3198: "Terminology for Policy-Based Management".

NOTE: Available at: http://www.ietf.org/rfc/rfc3198.txt

- IETF RFC 3261: "SIP: Session Initiation Protocol".

NOTE: Available at http://www.ietf.org/rfc/rfc3261.txt

- IETF RFC 3272: " Overview and Principles of Internet Traffic Engineering".

NOTE: Available at http://www.ietf.org/rfc/rfc3272.txt

- IETF RFC 3583: "Requirements of a Quality of Service (QoS)Solution for Mobile IP".

NOTE: Available at: http://www.ietf.org/rfc/rfc3583.txt

- IETF RFC 3726: "Requirements for Signaling Protocols".

NOTE: Available at: http://www.ietf.org/rfc/rfc3726.txt

- IETF RFC 4080: "Next Steps in Signaling (NSIS): Framework".

NOTE: Available at http://www.ietf.org/rfc/rfc4080.txt

- IETF RFC 4094: "Analysis of Existing Quality of Service Signaling Protocols".

NOTE: Available at: http://www.ietf.org/rfc/rfc4094.txt

- Cisco: "service provider Quality of Service".

NOTE: Available at www.cisco.com/warp/public/cc/so/neso/sqso/spqos_wp.pdf

- IEEE Communication Surveys 2005, Vol. 7, issue.1: "On The Building Blocks Of Quality Of Service In Heterogeneous IP Networks".

- IST TEQUILA project.

NOTE: Available at: http://www.ist-tequila.org/

- IST MUSE Project.

NOTE: Available at www.ist-muse.org

- ICN 2005: 4th International Conference on Networking (April 17-21, 2005) Proceedings: "Quality of Service Solutions in Satellite Communication" (ESA TRANSAT project).

- DSL Forum, DSL Evolution, TR 059: "Architecture Requirements for the Support of QoS-Enabled IP Services".

NOTE: Available at http://www.dslforum.org/techwork/treports.shtml

- API: "Generic QoS (GQoS) API; part of the Windows Sockets 2.0 (Winsock 2)".

NOTE: Available at
http://www.microsoft.com/technet/prodtechnol/windows2000serv/maintain/featusability/qoscomp.mspx

- The Open Group: "Resource Reservation Setup Protocol API (RAPI), C809" .

NOTE: Available at: http://www.opengroup.org/pubs/catalog.saved/c809.htm

- IEEE Transactions on communications, Vol. 35, issue 4, p 435-438 (April 1987): "On Packet Switches with Infinite Storage".

- ACM SIGCOMM Computer Communication Review, Symposium proceedings on Communications architectures and protocols SIGCOMM'88, Vol. 18, issue 4, p 314-329 (August 1988): "Congestion Avoidance and Control".

- IEEE/ACM Transactions on Networking, Vol. 1, issue 4, pp. 397-413 (August 1993): "Random Early Detection Gateways for Congestion Avoidance".

- IEEE Network, Special Issue on Transmission and Distribution of Digital Video, Vol. 12, issue 6, pp. 64-79 (November/December 1998): "An Overview of Quality-of-Service Routing for the Next Generation High-Speed Networks: Problems and Solutions".

History

Document history		
V1.1.1	December 2006	Publication