INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

Y.1291

(05/2004)

SERIES Y: GLOBAL INFORMATION INFRASTRUCTURE, INTERNET PROTOCOL ASPECTS AND NEXT GENERATION NETWORKS

Internet protocol aspects – Architecture, access, network capabilities and resource management

An architectural framework for support of Quality of Service in packet networks

ITU-T Recommendation Y.1291

ITU-T Y-SERIES RECOMMENDATIONS

GLOBAL INFORMATION INFRASTRUCTURE, INTERNET PROTOCOL ASPECTS AND NEXT GENERATION NETWORKS

GLOBAL INFORMATION INFRASTRUCTURE	
General	Y.100–Y.199
Services, applications and middleware	Y.200–Y.299
Network aspects	Y.300–Y.399
Interfaces and protocols	Y.400–Y.499
Numbering, addressing and naming	Y.500–Y.599
Operation, administration and maintenance	Y.600–Y.699
Security	Y.700–Y.799
Performances	Y.800–Y.899
INTERNET PROTOCOL ASPECTS	
General	Y.1000–Y.1099
Services and applications	Y.1100–Y.1199
Architecture, access, network capabilities and resource management	**Y.1200–Y.1299**
Transport	Y.1300–Y.1399
Interworking	Y.1400–Y.1499
Quality of service and network performance	Y.1500–Y.1599
Signalling	Y.1600–Y.1699
Operation, administration and maintenance	Y.1700–Y.1799
Charging	Y.1800–Y.1899
NEXT GENERATION NETWORKS	
Frameworks and functional architecture models	Y.2000–Y.2099
Quality of Service and performance	Y.2100–Y.2199
Service aspects: Service capabilities and service architecture	Y.2200–Y.2249
Service aspects: Interoperability of services and networks in NGN	Y.2250–Y.2299
Numbering, naming and addressing	Y.2300–Y.2399
Network management	Y.2400–Y.2499
Network control architectures and protocols	Y.2500–Y.2599
Security	Y.2700–Y.2799
Generalized mobility	Y.2800–Y.2899

For further details, please refer to the list of ITU-T Recommendations.

ITU-T Recommendation Y.1291

An architectural framework for support of Quality of Service in packet networks

Summary

This Recommendation provides an architectural framework for support of Quality of Service (QoS) in packet networks. Central to the architectural framework is a set of generic network mechanisms (or QoS building blocks) for controlling the network service response to a service request, which can be specific to a network element, or for signalling between network elements, or for controlling and administering traffic across a network. Distributed across three logical planes (namely the Control Plane, Data Plane and Management Plane), the building blocks can be used in combination to form various approaches for delivering the satisfactory collective effect of varying service performance required by a range of applications, such as file transfer and multimedia conferencing.

Source

ITU-T Recommendation Y.1291 was approved on 7 May 2004 by ITU-T Study Group 13 (2001-2004) under the ITU-T Recommendation A.8 procedure.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure e.g. interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementors are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database.

CONTENTS

Page

ITU-T Recommendation Y.1291

An architectural framework for support of Quality of Service in packet networks

1 Scope

This Recommendation provides an architectural framework for support of Quality of Service (QoS) in packet networks. Central to the architectural framework is a set of QoS building blocks distributed across three logical planes (namely the Control Plane, Data Plane and Management Plane) to control network performance, even in case of network resource contention. Ultimately the building blocks are to help deliver "the collective effect of service performance which determines the degree of satisfaction of a user of the service".

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

- ITU-T Recommendation E.360.1 (2002), *Framework for QoS routing and related traffic engineering methods for IP-, ATM- and TDM-based multiservice networks*.
- ITU-T Recommendation E.360.2 (2002), *QoS routing and related traffic engineering methods – Call routing and connection routing methods*.
- ITU-T Recommendation E.360.3 (2002), *QoS routing and related traffic engineering methods – QoS resource management methods*.
- ITU-T Recommendation E.360.4 (2002), *QoS routing and related traffic engineering methods – Routing table management methods and requirements*.
- ITU-T Recommendation E.360.5 (2002), *QoS routing and related traffic engineering methods – Transport routing methods*.
- ITU-T Recommendation E.360.6 (2002), *QoS routing and related traffic engineering methods – Capacity management methods*.
- ITU-T Recommendation E.360.7 (2002), *QoS routing and related traffic engineering methods – Traffic engineering operational requirements*.
- ITU-T Recommendation E.361 (2003), *QoS routing support for interworking of QoS service classes across routing technologies*.
- ITU-T Recommendation E.860 (2002), *Framework of a service level agreement*.
- ITU-T Recommendation G.114 (2003), *One-way transmission time*.
- ITU-T Recommendation G.1000 (2001), *Communications Quality of Service: A framework and definitions*.
- ITU-T Recommendation G.1010 (2001), *End-user multimedia QoS categories*.
- ITU-T Recommendation I.350 (1993), *General aspects of quality of service and network performance in digital networks, including ISDNs*.

- ITU-T Recommendation J.112 (1998), *Transmission systems for interactive cable television services*.
- ITU-T Recommendation J.162 (2004), *Network call signalling protocol for the delivery of time-critical services over cable television networks using cable modems*.
- ITU-T Recommendation J.163 (2004), *Dynamic quality of service for the provision of real-time services over cable television networks using cable modems*.
- ITU-T Recommendation J.170 (2002), *IPCablecom security specification*.
- ITU-T Recommendation J.174 (2002), *IPCablecom interdomain quality of service*.
- ITU-R Recommendation M.1079-2 (2003), *Performance and quality of service requirements for International Mobile Telecommunications-2000 (IMT-2000) access networks*.
- ITU-T Recommendation X.805 (2003), *Security architecture for systems providing end-to-end communications*.
- ITU-T Recommendation Y.1221 (2002), *Traffic control and congestion control in IP-based networks*.
- ITU-T Recommendation Y.1540 (2002), *Internet protocol data communication service – IP packet transfer and availability performance parameters*.
- ITU-T Recommendation Y.1541 (2002), *Network performance objectives for IP-based services*.

3 Definitions

This Recommendation does not define new terms.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations:

DiffServ	Differentiated Services
DQoS	Dynamic QoS
IETF	Internet Engineering Task Force
IntServ	Integrated Services
ITU-T	International Telecommunication Union – Telecommunication Standardization Sector
LSP	Label Switched Path
MPLS	Multi-Protocol Label Switching
MTA	Multimedia Terminal Adaptor
QoS	Quality of Service
RSVP	Resource ReSerVation Protocol
SLA	Service Level Agreement

5 Introduction

Quality of Service (QoS) ultimately is about supporting the characteristics and properties of specific applications. Yet different applications may have quite different needs. For example, for telemedicine, the accuracy of the delivery is more important than overall delay or packet delay variation (i.e., jitter), while for IP telephony, jitter and delay are key and must be minimized. A number of ITU-T Recommendations deal with QoS. ITU-T Rec. E.800 defines QoS as "the collective effect of *service performance* which determines the degree of satisfaction of a user of the service". Given that ITU-T Rec. E.800 considers support, operability, serviceability and security all part of service performance, this QoS definition is comprehensive in scope. Expanding on the E.800 QoS concept, ITU-T Rec. G.1000 breaks down *service performance* (or service quality) into functional components and links it to network performance such as defined in ITU-T Recs I.350, Y.1540 and Y.1541. Complementary to ITU-T Rec. G.1000, which gives a framework, ITU-T Rec. G.1010 provides end-user-centric application requirements in terms of broad categories (such as interactive, error tolerant). Concerning specific applications or performance parameters, among related standards, ITU-R Rec. M.1079-2 defines the end-to-end speech and data quality and performance requirements for IMT-2000 access networks, while ITU-T Rec. G.114 specifies the bounds for transmission time for connections across a digital network.

To deliver required network performance, certain mechanisms need to be in place within the network. These network mechanisms are to control and deliver various network service responses, even in case of network resource contention. IETF RFC 2990 summarizes the possible characteristics of the controlled service response to a specific service request: *consistent and predictable*, *at a level equal to or above a guaranteed minimum*, or *established in advance.* For example, in case of network resource contention or congestion, to maintain the expected service response requires a variety of means working at different time-scales, from those for careful network planning based on traffic patterns over a long period to those for differential resource allocation and admission control based on the current network load condition. These and other mechanisms (e.g., a signalling method for indicating the desired level of network performance) are the focus of the architectural framework for QoS support. In particular, this Recommendation identifies a set of generic QoS network mechanisms and provides a structure for them. Ultimately the network mechanisms are to be used in combination to deliver the satisfactory collective effect of varying service performance required by a wide range of applications. The application-independent aspect of the identified architectural framework distinguishes it from application-specific QoS architectures such as defined in ITU-T Rec. H.360, which is specific to multimedia applications.

6 QoS building blocks

Key to the QoS architectural framework is a set of generic network mechanisms for controlling the network service response to a service request, which can be specific to a network element, or for signalling between network elements, or for controlling and administering traffic across a network. (Please note that the building blocks should not be deemed end to end.) As depicted in Figure 1, the building blocks are organized into three planes:

– Control Plane, which contains mechanisms dealing with the pathways through which user traffic travels. These mechanisms include admission control, QoS routing, and resource reservation.

– Data Plane, which contains mechanisms dealing with the user traffic directly. These mechanisms include buffer management, congestion avoidance, packet marking, queuing and scheduling, traffic classification, traffic policing and traffic shaping.

– Management Plane, which contains mechanisms dealing with the operation, administration, management aspects of the network. These mechanisms include Service Level Agreement (SLA), traffic restoration, metering and recording, and policy.

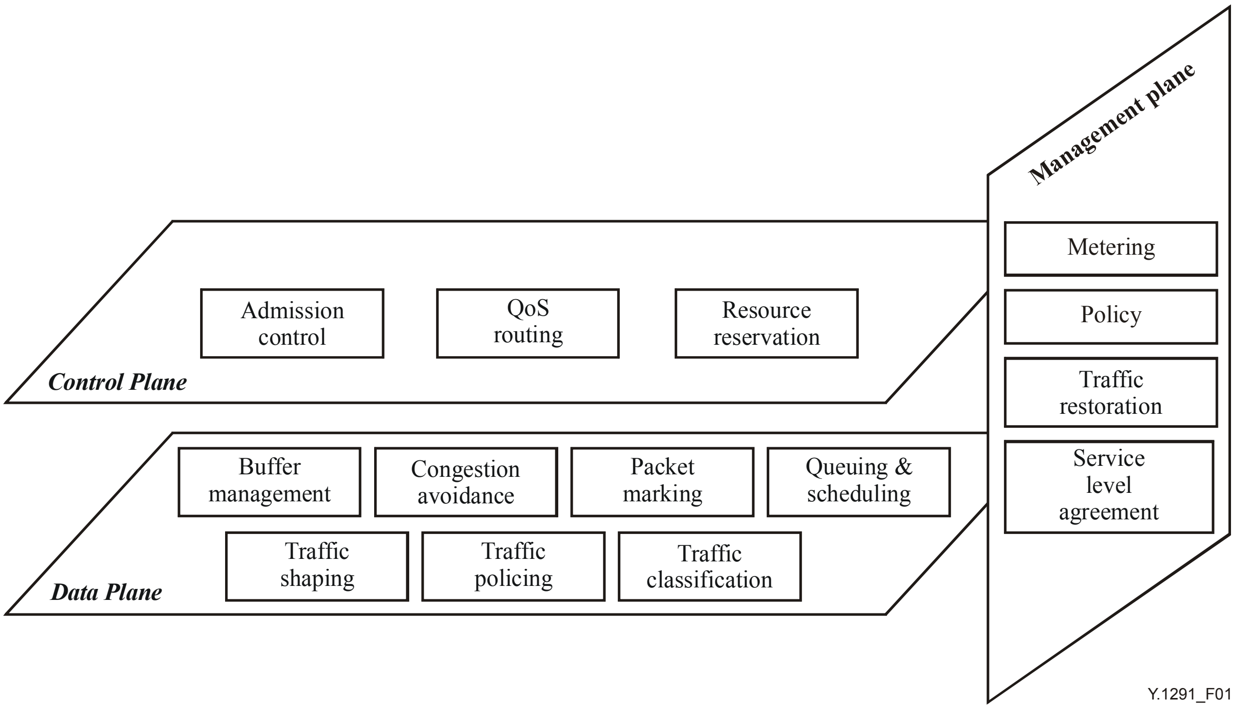


Figure 1/Y.1291 – Architectural framework for QoS support

A QoS building block may be specific to a network node (as exemplified by buffer management) or applicable to a network segment (as exemplified by QoS routing). The latter, in particular, requires signalling between network nodes, whether they are part of a network segment that is end to end, end to edge, edge to edge, or network to network. Signalling can take place in any of the three logical planes. When taking place in the Control or Management Plane, signalling entails the use of a signalling protocol. Because of its unique properties, this Recommendation treats signalling as part of interactions among QoS building blocks and discusses it in the corresponding clause.

It is important to note that the QoS architectural framework is logical and puts no constraint on how a building block is realized. As such, the implementation of a building block can be, for instance, distributed or centralized. The following clauses further describe the building blocks according to the planes.

7 Control-plane mechanisms

7.1 Admission control

This mechanism controls the traffic to be admitted into the network. Normally the admission criteria are policy driven [IETF RFC 2753]. 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. 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.

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. Appendix II summarizes an experimental measurement-based QoS approach. 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.

7.2 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)[1]. 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 [Chen]. 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 [IETF RFC 2386].

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 [IETF RFC 2205] can be used for conveying the flow's requirements and characteristics and OSPF extensions as defined in IETF RFC 2676 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 [Apostolopoulos].

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

[1] 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.

account a variety of constraints such as traffic attributes, network constraints, and policy constraints [IETF RFC 3272]. 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 E.360.x series of Recommendations 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 Rec. E.361 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 Rec. E.361 prescribes means for signalling QoS routing parameters across networks employing different routing technologies.

7.3 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 [IETF RFC 2205].

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.

8 Data-plane mechanisms

8.1 Queue (or buffer) management

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 [IETF RFC 2309]. 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.

A common criterion for dropping packets is the queue reaching the maximum size. Packets are dropped when the queue is full. What packets drop depends 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) [Floyd], 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 favoring 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.

8.2 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 [1987] 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) [Jacobson, 1988]. 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. IETF RFC 3168 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.

8.3 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.

8.4 Packet marking

Packets can be marked according to the specific service classes that they will receive in the network on a per-packet basis. Typically performed by an edge node, packet marking involves assigning a value to a designated header field of a packet in a standard way. (For example, the type of service in the IP header or the EXP bits of the MPLS shim header [IETF RFC 3032] is used to codify externally observable behaviours of routers in the *DiffServ* [IETF RFC 2474] or *MPLS-DiffServ* [IETF RFC 3270] approach.) 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.

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 (IETF RFC 2748) 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 [IETF RFC 2996].

8.5 Traffic classification

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

8.6 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.

8.7 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.

9 Management-plane mechanisms

9.1 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) [IETF RFC 3198], 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 Rec. Y.1540 or technology specific such as the performance and traffic parameters used in *IntServ* or *DiffServ*. Overall, ITU-T Rec. E.860 defines a general SLA framework for a multi-vendor environment.

9.2 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.

9.3 Traffic restoration

Restoration is broadly defined here as the mitigating response from a network under conditions of failure and should be considered at multiple layers. At the bottom of the layered stack, optical networks are now capable of providing dynamic ring and mesh protection and restoration functionality at the wavelength level. At the SONET/SDH layer reliability capability is provided with Automatic Protection Switching (APS) as well as self-healing ring and mesh architectures. ATM provides similar capabilities. Re-routing is traditionally used at the IP layer to restore service following link and node failures and can be end-to-end or local (fast reroute). Re-routing at the IP layer occurs after a period of routing convergence, which may require seconds to minutes to complete. MPLS now provides recovery at the IP layer prior to convergence.

There are two types of network failures:

- Node Failure: Failure of a network element (e.g., router card) in a network node or office. This type of failure is typically dealt with by designing redundancy features in network elements to minimize failure impact. Catastrophic failures such as power outages and natural disasters however may take down an entire network node. In which case, through traffic can be re-routed over spare links designed around the failed node.
- Transport Link Failure: Failure of a link (e.g., T1, OC-3) connecting two network nodes. Typically links can fail due to link element failure (e.g., line card) (which can then take down a single link) or, more seriously, a fibre cut (which can then disrupt a large number of links). Service providers can design additional spare capacity to mitigate the impact of such failures and restore traffic flows until the failure is repaired.

Note that some of these terms are generally layer specific and one should consider carefully the multiple layers involved in the overall design. For example a link failure at the physical layer might impact many links and paths at the IP layer.

As in the case of admission control, certain traffic streams related to critical services may require higher restoration priority than others. A service provider needs to plan for adequate levels of spare resources such that QoS SLAs are in compliance under conditions of restoration. Typical parameters for measuring service restorability are time-to-restore and the percentage of service restorability. The details of the priority levels can be found in Annex A.

9.4 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. IETF RFC 2748, 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).

10 Interactions among building blocks

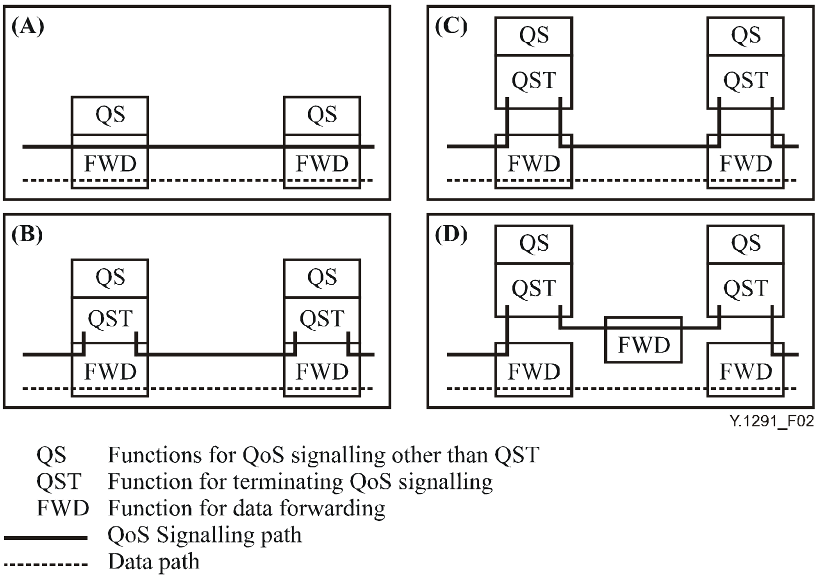
A comprehensive QoS solution typically employs multiple building blocks across the Control Plane, Data Plane and Management Plane. QoS parameters therefore need to be exchanged between the various building blocks. These parameters include transaction performance at the packet level (e.g., delay and packet 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.

10.1 QoS signalling

QoS signalling is mainly for conveying application (or network) performance requirements, reserving network resources across the network, or discovering QoS routes. Depending on whether the signalling information is part of the associated data traffic, QoS signalling may be effected in or out of band:

– In band: The QoS signal is part of the associated data traffic, typically presented in a particular header field (e.g., the TOS field in IPv4 as in *DiffServ* and 802.1p) of the data packets. Taking place in the Data Plane, in-band signalling neither introduces additional traffic into the network nor incurs setup delay for the data traffic. Naturally such a type of signalling is not suitable for resource reservation or QoS routing, which needs to be done *a priori* before data transmission.

– Out of band: The QoS signal, being carried by dedicated packets, is separated from the associated data traffic. In addition, QoS signalling may be hop-by-hop or end-to-end. In the hop-by-hop case (shown as Case B in Figure 2), the signalling information is likely to be modified by intermediary nodes. In contrast, in the end-to-end case (shown as Case A in Figure 2), the signalling information is not modified by intermediary nodes. As a result, out-of-band signalling introduces extra traffic into the network and incurs an overhead for delivering desired network performance. In addition, it entails the use of a signalling protocol and further processing above the network layer, which tends to render not as fast responses as in-band signalling. Nevertheless out-of-band signalling lends itself naturally for resource reservation or QoS routing.



QS	Functions for QoS signalling other than QST
QST	Function for terminating QoS signalling
FWD	Function for data forwarding
———	QoS Signalling path
----------	Data path

Figure 2/Y.1291 – Illustration of different forms of QoS signalling

Similarly, depending on whether the signalling path is closely tied to the associated data path, QoS signalling may be viewed as path coupled or decoupled:

– Path-coupled: QoS signalling messages are routed only through the nodes that are potentially on the data path. As such, in-band signalling by definition is path-coupled but out-of-band signalling may or may not be the case. Path-coupled signalling implies that signalling nodes must be co-located with routers. Such an arrangement has, on one hand, the advantage of reduced overall signalling processing cost (since it leverages network-layer routing tasks) but, on the other hand, the disadvantage of inflexibility in upgrading routers or in integrating control entities (e.g., policy servers) not on the data path (or non-traditional routing methods). If a path-coupled mechanism involves a signalling protocol, this means routers need to support the protocol and be able to process related signalling messages. An example of a path-coupled signalling protocol is RSVP.

– Path-decoupled: QoS signalling messages are routed through nodes that are not assumed to be on the data path. As such, only out-of-band signalling may be path-decoupled. Path-decoupled signalling implies that the entity terminating QoS signalling should be dedicated and separated from the forwarding entity, which is normally located in routers. In contrast to path-coupled signalling, it has the advantage of flexibility in deploying and upgrading signalling nodes independent of routers or in integrating control entities not on the data path but has the disadvantage of added complexity and cost in overall processing and operational tasks. Case C and Case D in Figure 2 further illustrate path-decoupled signalling.

10.2 Intra-plane

This subject is for further study.

10.3 Inter-plane

Control and data plane

Mapping of Y.1541 QoS Class to DSCP

An association of ITU-T Rec. Y.1541 QoS classes with Diffserv Per Domain Behaviours (PDBs) is presented in Appendix VI/Y.1541:

- PDB based on Expedited Forwarding PHB: Y.1541 Classes 0 and 1.
- PDB based on Assured Forwarding PHB: Y.1541 Classes 2, 3, and 4.
- PDB based on Best Effort (Default) PHB: Y.1541 Class 5.

11 Security considerations

In general, ITU-T Rec. X.805 provides a network-security architecture useful for examining the security properties of and devising safeguards against QoS building blocks and solutions.

11.1 Data plane

On the data plane traffic is typically processed according to the packet header information. Packets can be marked by assigning a value to a designated header field or classified based on multi-fields of the packet header (such as the IP five tuples). Traffic shaping, policing and queuing can then be done on the basis of the packet classification and marking. As such, integrity of the packet headers is essential for validity and security of a QoS approach. Malicious modification and fabrication of the packet header information must be prevented.

It is also important to note that although a packet mark can be done or changed by a host or any network nodes, the mark done by an edge node is desirable. An edge node in general has a trust relationship with core nodes. So if done by a host, the packet mark should be checked and may be changed as necessary by an edge node.

11.2 Management and control plane

The control plane and the management plane deal with traffic at the flow or aggregate level. A flow is also identified and described by, e.g., the IP five tuples or a MPLS label in packet header, which is constant during the lifecycle of the flow.

Admission control performed at the edge nodes is helpful against masquerading attacks and the resulting congestion by unauthorized traffic. The edge nodes may be trusted by the core nodes and may have a view of the overall network resource utilization. Whether performed in a centralized or distributed way, admission control should cover authentication and authorization.

Resource reservation is closely tied to admission control. A reservation request can be initiated by an end host or a service-supporting node situated in the network. Malicious resource requests may result in illegal excessive reservation, resource exhaustion and denial of service. Safeguards to prevent such malicious requests are desirable.

In general, network security mechanisms such as firewalls and intrusion detection can help protect the network interfaces, whether QoS is involved or not. Also entities responsible for authentication should have safeguards against DoS attacks.

11.3 QoS Signalling

To protect against interception, modification and fabrication attacks, the QoS signalling should make use of authentication and integrity mechanisms, such as RIPEMD160 or SHA-1 (Secure Hash Algorithm 1). The use of security mechanisms may have performance implications. Because signalling traffic is normally much less than data traffic, the network performance impacts due to secure out-of-band (or path-decoupled) signalling should be less than secure in-band (or path-coupled) signalling. In addition entities responsible for signalling should have safeguards against DoS attacks.

12 Example approaches

To illustrate how QoS building blocks interact and form various QoS approaches, this clause describes four standardized approaches: integrated services (*IntServ*), differentiated services (*DiffServ*), Multi-Protocol Label Switching (MPLS), and IPCablecom Dynamic QoS. (Note that IETF RFC 2998 integrates the IntServ and DiffServ approaches.) Since other more comprehensive approaches are just emerging and of evolutionary nature, examples are shown in Appendices I and II.

12.1 IntServ

Primarily for supporting real-time delay sensitive applications, the *IntServ* (see, e.g., [IETF RFC 1633]) approach is built on the understanding that a flow serviced at a rate slightly higher than its data rate has a bounded delay and the network can guarantee the delay bound of a flow by per-flow resource reservation. With this approach, an application, before sending data, first signals to the network the desired service request, including specifics such as its traffic profile and bandwidth and delay requirements. The network then determines whether it can allocate adequate resources (e.g., bandwidth or buffer space) to deliver the desired performance of the service request. Only after the request is granted can the application start to send data. As long as the application honours its traffic profile, the network meets its service commitment by maintaining per-flow state and by using advanced queuing disciplines (e.g., weighted fair queuing) for link sharing. The building blocks relevant to the *IntServ* approach include admission control, queuing, resource reservation, traffic classification, and traffic policing. In particular, the signalling protocol RSVP is used to reserve resources. The network may accept or reject a reservation request via admission control based on resource availability. A successful reservation request results in installation of states at the RSVP-aware nodes. The building blocks interact by having access to the state information and other provisioned (thus relatively static) data objects.

12.2 DiffServ

The concept behind the *DiffServ* approach is treating a packet based on its class of service as encoded in its IP header. The service provider establishes with each user a service level agreement (or service level specification), which, among other things, specifies how much traffic a user may send within any given class of service. The ensuing traffic is classified (on a per-packet basis) into one of a small number of aggregated flows or classes and policed at the border of the service provider's network. Once the traffic enters the network, routers provide it with differentiated treatment. In contrast to the *IntServ* approach, the treatment is based not on a per-flow basis, but solely on the indicated class of service. The overall network is set up so as to meet all service level agreements. The relevant building blocks (which include buffer management, packet marking, service level agreement, traffic metering and recording, traffic policing, traffic shaping, and scheduling) interact with each other in a relatively static way, primarily through provisioned data objects.

12.3 MPLS

Initially developed for the purpose of interworking between the IP and ATM (or Frame Relay) networks, MPLS [IETF RFC 3031] achieves substantial gains in packet forwarding speed through the use of short, layer-2-like labels. Upon entering the MPLS network, a packet is assigned once and for all a Forward Equivalence Class (FEC), which is encoded as a fixed length string known as a label. When the packet is forwarded to the next hop, the label is sent along with it. At the next hop, the label is used as an index into a pre-configured table to identify the following hop, and a new label. The old label is replaced with the new label and the packet is forwarded to the following hop. The process continues till the packet reaches the destination. In other words, packet forwarding in MPLS is entirely label driven, whereby packets assigned the same FEC are forwarded the same way. Furthermore, labels are meaningful only to the pair of routers sharing a link, and only in one direction – from a sender to the receiver. The receiver, however, chooses the label and negotiates its semantics with the sender by means of a label distribution protocol. MPLS in its basic form is particularly useful for traffic engineering. To provide explicit QoS support, MPLS makes use of certain elements in the *IntServ* and *DiffServ* approaches. The label distribution protocol, for example, can be based on a resource reservation protocol [IETF RFC 3209]. With it, required network resources along a label switched path can thus be reserved during its set-up phase to guarantee the QoS of packets passing through the path. In addition, by using the label and certain EXP bits of the shim header that carries the label to represent the differentiated service classes, packets of the same FEC can be subject to *DiffServ* treatment [IETF RFC 3270]. The relevant building blocks for MPLS include buffer management, packet marking, QoS routing, queuing, resource reservation, traffic classification and traffic shaping. They interact through the label-switched-path state information installed in each MPLS node by a label distribution protocol and through provisioned data objects.

12.4 IPCablecom dynamic QoS

To support interactive multimedia applications over the IPCablecom access network, ITU-T Rec. J.163 specifies an approach based on dynamic per flow resource reservation. The access network connects the Multimedia Terminal Adaptor (MTA) to the Access Node as defined in ITU-T Rec. J.112. Resources are allocated on the J.112 network for each individual flow associated with an application session, per subscriber, on an authorized and authenticated basis.

Central to the dynamic QoS approach are the Dynamic QoS (DQoS) Gates and Gate Controller. Using the Common Open Policy Service Protocol (COPS) per IETF RFC 2748, the gate controller controls the existence and operation of the gates.

DQoS gates are implemented on the access node between the J.112 network and an IP backbone using the J.112 packet classification and filtering functions. Unidirectional in nature, a DQoS gate is a logical entity associated with a session. If a gate is "closed", data in transit in the J.112 access network may either be dropped or receive simply the best-effort service, depending on the provider policy.

The gate controller is implemented on the Call Management Server, which normally manages multimedia sessions initiated by MTAs through the network-controlled call signalling (as defined in ITU-T Rec. J.162) or distributed call signalling (as defined in IETF RFC 3261). The controller is responsible for the policy decision on whether to create as well as open a gate. Opening a gate involves admission control upon receipt of a resource management request (by way of RSVP) and resource reservation as necessary in the network. Worth noting is that resource reservation is done in two phases. At the end of the first phase, resources are reserved but are not yet available to the MTAs. Only at the end of the second phase, are the gates at the ANs opened and resources made available to the MTAs. The reserve and commit model ensures that resources are available before signalling to the terminating party that a session is being initiated and resources are committed only when they are required.

The relevant building blocks for the IPCablecom DQoS approach include mainly admission control, queuing, resource reservation, traffic classification, traffic policing and policy. The signalling protocols RSVP and COPS are used to reserve and commit resources. The network may accept or reject a reservation request via admission control based on resource availability or policy. A successful reservation request results in installation of states at the RSVP-aware nodes. The building blocks interact by having access to the state information and other provisioned data objects.

Annex A

Traffic priority levels

Quality of Service (QoS) expectations for services in packet networks can be considered from two perspectives. Transaction packet performance objectives (e.g., packet loss and delay) are governed by the transaction classes specified in ITU-T Recommendations such as Y.1541 for IP services and I.356 for ATM services. These classes cover a wide range of services including voice, data, and multimedia applications. The associated parameters define acceptable performance levels (e.g., lost packets) for each transaction class. Reliability expectations expressed as a priority relate to the setup of the link or "connection" such as a Multi-Protocol Label Switching (MPLS) Label Switched Path (LSP) over which a packet transaction can be routed in the network. Mechanisms used to achieve these QoS objectives include call and connection routing methods and QoS resource allocation methods such as bandwidth allocation, priority routing, priority queuing, and transport restoration. The scope of this annex is the reliability of such LSPs, expressed in the form of a priority, and the need for specifying priority levels for QoS signalling.

Traffic priorities play an important role in providing acceptable service reliability/availability to customers in communications networks. For example, emergency communications require the highest available admission control priority under conditions involving natural disasters or terrorist attacks. In today's Public Switched Telephone Networks (PSTN) this priority level is unique. The desired reliability/availability can be requested as a priority level for a particular network function that, in turn, determines the setup of an LSP. Two network functions for priority consideration in evolving packet networks are:

- Connection Admission Control: Admission control policies give preference to traffic streams deemed to be more critical by a service provider (e.g., emergency communications) under conditions of congestion. Admission control priority is a way of giving preference to admit higher priority LSPs ahead of lower priority LSPs.
- Restoration: Restoration is broadly defined here as the mitigating response from a network under conditions of failure. Potential methods for failure recovery include Automatic Protection Switching for line/path protections and shared mesh restoration methods. Critical service traffic streams can request restoration with higher priority. Such a traffic stream can then be routed over an LSP that has the appropriately "marked" restoration priority level.

The establishment of traffic priorities should permit maximum flexibility for implementation from the perspective of service providers. The priority levels should meet the following requirements:

- The total number of priority classes should be small in number in order to ensure scalability.
- Sub-divisions within any priority class should be avoided in order to ensure simplicity.
- The priority levels are relative and not associated with specific parameters (e.g., time to restore) and their values.
- Service providers should be allowed to choose the number of priority levels from the available set for their service offerings. Accordingly, they can construct Service Level Agreements (SLA) for any given priority class treatment to their customers, including other service providers (network-network interface).

For customer service traffic, four priority levels are identified for connection admission control:

- Critical: Unique priority level reserved for emergency communications traffic for all service providers, domestic and international.
- High: Service examples include other government services, key business customers, virtual private networks.
- Normal: Service examples include residential voice services.
- Best Effort: Service examples include Internet Service Provider services.

For restoration three priority levels are identified: High, Normal, and Best Effort. The service examples listed above apply here; emergency communications would request High Priority.

As mentioned above, a service provider may offer specific priority service offerings based on available network capabilities and customer needs. For example, a service provider may choose to offer services with all four of the defined connection admission priorities but with only High and Normal restoration priorities. A pure ISP provider on the other hand may choose to offer only the Critical and Best Effort connection admission priorities and the Best Effort restoration priority.

Appendix I

A comprehensive QoS approach based on independent resource control

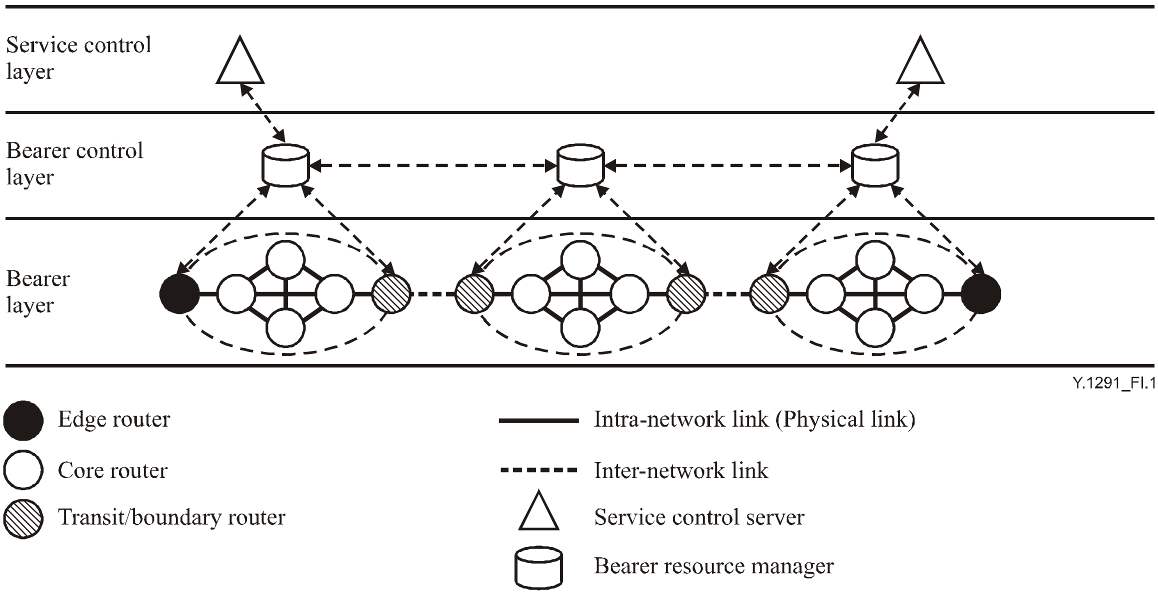


Figure I.1/Y.1291 – Comprehensive QoS approach based on independent resource control

For supporting services of varying performance requirements over a single IP core network and guaranteeing QoS of connection-oriented and real-time services (such as IP telephony), a comprehensive QoS approach based on independent resource control is developed as depicted in Figure I.1. The approach integrates MPLS, DiffServ, traffic engineering and policy management.

Services requiring QoS guarantee are categorized according to general service types (e.g., voice) or QoS treatment levels (e.g., EF). For manageability and stability of the network, IP core network of a network provider is divided into multiple administrative domains. Such division is flexible and may not be the same as that of routing domains. For example, an administrative domain may be as small as to contain only one edge router, or as large as to contain a whole operator network.

A bearer resource manager (BRM) is an independent resource control function that manages all bearer resources over each administrative domain and could be implemented in one or multiple boxes. The BRM records and maintains the network topology and resource database (NTRD). Based on the NTRDs, the BRM makes intra-domain path selection, resource allocation and admission control for a service flow. The BRMs of different domains interact through signalling to perform resource control for inter-domain application flows. In addition, a BRM may also have functions like policy management, SLA management, LSP traffic metering, and interface with AAA servers.

A variety of service control servers (SCS) are responsible for controlling various service requests (e.g., voice call signaling), identifying the originating and terminating point of each service request, translating number (or name) into IP address, and then sending the resource requests to the BRM of the originating domain. For services with QoS requirements but without service control servers, like point-to-point services, hosts can initiate a QoS service request through RSVP or other QoS signalling protocols. Here, RSVP is only for hosts to request QoS guarantee and routers need not support RSVP for per-flow resource reservation. The equipment deployed to process QoS service requests of hosts can be viewed as a particular kind of SCS.

A BRM receives resource requests from the SCS within its administrative domain or from other BRM. It processes them and then notifies the responses back to the SCS. At the same time, if a service flow resource request is admitted, the BRM notifies the flow identification, path and QoS attributes to the ingress edge routers. The ingress edge router identifies, classifies, marks, policies, shapes, and encapsulates the packets of a flow with the QoS information specified by the BRM.

For the service flows travelling across multiple network providers, generally there are application gateways and boundary routers between different network providers which interlink through the fixed link resources and the specified inter-network SLAs. Different network providers may deploy different QoS mechanisms in their networks. In this case, BRMs only manage the intra-network link resources, whereas application gateways or boundary routers manage the inter-network link resources by the specified SLAs and an application gateway or boundary router acts as the ingress or egress edge router.

The relevant building blocks for this approach almost involve all blocks illustrated in Figure 1. BRM serves as a physically independent control and management plane. The building blocks interact primarily through signalling at a per-flow level and on the basis of per-LBN resource management. There is a clear signalling interface between control plane and data plane.

I.1 Implementation flexibility for packet networks with MPLS support

In this case, it is assumed that DiffServ-aware MPLS is supported in IP core networks.

MPLS LSP technology is deployed to pre-provision a logical bearer network (LBN) for each service class over the underlying IP network manually or automatically through RSVP-TE or CR-LDP protocol. For service flows belonging to a service class, path selection, resource allocation, admission control and label forwarding are dealt within the same one LBN. The topology planning and bandwidth reservation of each LBN depends on traffic metering and forecasting data, administrative policies and SLA, which can be adjusted automatically or manually for LSP protections, capacity changes or network performance optimization in accordance with traffic engineering constraints.

Within the remaining resource of the underlying packet networks, BE traffic without QoS requirements are still routed and forwarded by conventional IP routing and forwarding methods with or without DiffServ.

The BRM records and maintains a network topology and resource database (NTRD) separately for each LBN. Based on the NTRDs and policies, the BRM makes intra-domain path selection, resource allocation and admission control for a service flow within its corresponding LBN. As for the remaining resource of the underlying packet networks, the BRM could also perform resource allocation and admission control.

The QoS path information for a flow specified by BRM is a multi-layer label stack that represents a concatenated LSP set. The edge router encapsulates the packets with this label stack, which in turn makes the intermediate transit routers forward the packets of a flow along the specified path in terms of the label stack and the specified priority.

I.2 Implementation flexibility for packet networks without MPLS support

In this case, admission control and resource reservation are dynamically applied with the link-by-link resource reservation, and MPLS capability is not required to the bearer layer. Routing and forwarding of all traffic is under the control of traditional IP routing protocols and IP Diffserv.

BRM is deployed to directly manage all of the physical link resources within each administrative domain. The BRM holds and maintains a network topology and resource database (NTRD). Based on the information in the NTRD, the BRM handles route look-up, link-by-link resource reservation and admission control for each flow that requires QoS guarantee. If a flow is admitted with high priority, it will not interfere with other traffic flows.

I.3 Implementation flexibility for distributed resource control

In this case, LBNs are virtual links (called QoS pipes) between ingress-egress ER pairs in a network domain. A QoS pipe is set up to carry aggregated flows of a specific service or QoS class.

If the BRM function is implemented in edge routers (ER), per-flow resource control is distributed to the edges. The resource control function (RCF) on ER maintains the resource status table of the corresponding QoS pipes and accordingly performs admission control and resource allocation. It also processes QoS signalling.

QoS pipes are adjusted in a middle term or long term manually or automatically, which can be implemented by network management system.

Appendix II

Priority promotion scheme

The Priority Promotion Scheme (PPS) is a new scheme for traffic control that is still at the experimental stage. In a nutshell, PPS makes use of a form of admission control to achieve end-to-end QoS in a packet-based network. The main applications for such a scheme are interactive multimedia services such as voice over IP, video chat, and videoconferencing. Specifically, the scheme is based on end-to-end measurement of network resources by end systems. Before a session is established or even during a session, the source end system senses, measures, or probes the availability of network resources by sending out packets with priority one level lower than that of normal packets. The result is modification of the DiffServ Code Point (DSCP) value of the succeeding IP packets: the priority is raised or promoted to firmly establish the session, lowered to leave resources with existing sessions, or otherwise adjusted so that the number of packets does not exceed the available capacity. The network, i.e., output links of the routers or L2 switches is only assumed to support the per-class form of priority control that accompanies the DiffServ architecture. Having all end systems follow the above behaviour achieves end-to-end QoS without the maintenance of per-flow states in network nodes.

BIBLIOGRAPHY

[IETF RFC 1633] BRADEN (R.), *et al.*: Integrated Services in the Internet Architecture: an Overview, June 1994.

[IETF RFC 2205] BRADEN (R.), *et al.*: Resource ReSerVation Protocol (RSVP) – Version 1 Functional Specification, September 1997.

[IETF RFC 2309] BRADEN (R.), *et al.*: Recommendations on Queue Management and Congestion Avoidance in the Internet, April 1998.

[IETF RFC 2386] CRAWLEY (E.), *et al.*: A Framework for QoS-based Routing in the Internet, August 1998.

[IETF RFC 2474] NICHOLS (K.), *et al.*: Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers, December 1998.

[IETF RFC 2748] DURHAM (D.), *et al.*: The COPS (Common Open Policy Service) Protocol, January 2000.

[IETF RFC 2753] YAVATKAR (R.), *et al.*: A Framework for Policy-based Admission Control, January 2000.

[IETF RFC 2990] HUSTON (G.): Next Steps for the IP QoS Architecture, November 2000.

[IETF RFC 2996] BERNET (Y.): Format of the RSVP DCLASS Object, November 2000.

[IETF RFC 2998] BERNET (Y.), *et al.*: A Framework for Integrated Services Operation over Diffserv Networks, November 2000.

[IETF RFC 3031] ROSEN (E.), *et al.*: Multiprotocol Label Switching Architecture, January 2001.

[IETF RFC 3032] ROSEN (E.), *et al.*: MPLS Label Stack Encoding, January 2001.

[IETF RFC 3198] WESTERINEN (A.), *et al.*: Terminology for Policy-Based Management, November 2001.

[IETF RFC 3209] AWDUCHE (D.), *et al.*: RSVP-TE: Extensions to RSVP for LSP Tunnels, December 2001.

[IETF RFC 3261] ROSENBERG (J.), *et al.*: SIP: Session Initiation Protocol, June 2002.

[IETF RFC 3270] LE FAUCHEUR (F.), *et al.*: Multi-Protocol Label Switching (MPLS) Support of Differentiated Services, May 2002.

[IETF RFC 3272] AWDUCHE (D.): Overview and Principles of Internet Traffic Engineering, May 2002.

[Jacobson, 1988] JACOBSON (V.): Congestion Avoidance and Control*, Proceedings of ACM SIGCOMM'88*, pp. 314-329, August 1988.

[Lin *et al.*, 1997] LIN (D.), MORRIS (R.): Dynamics of Random Early Detection*, Proceedings of ACM SIGCOMM'97*, pp. 127-138, September 1997.

[Chen] CHEN (Shigang), NAHRSTEDT (Klara): An Overview of Quality-of-Service Routing for the Next Generation High-Speed Networks: Problems and Solutions*, IEEE Network, Special Issue on Transmission and Distribution of Digital Video*, Vol. 12, No. 6, pp. 64-79, November/December 1998.

[Apostolopoulos] APOSTOLOPOULOS (D.), *et al.*: Intra domain QoS Routing in IP Networks: A Feasibility and Cost Benefit Analysis, *IEEE Network*, Vol. 13, No. 5, pp. 42, September/October 1999.

[Floyd] FLOYD (S.), JACOBSON (V.): Random Early Detection Gateways for Congestion Avoidance, *IEEE/ACM Transactions on Networking*, Vol. 1, No. 4, pp. 397-413, August 1993.

[Nagle] NAGLE (J.): On Packet Switches with Infinite Storage. *IEEE Trans. on communications*, Vol. COM-35, pp. 435-438. April 1987.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series B	Means of expression: definitions, symbols, classification
Series C	General telecommunication statistics
Series D	General tariff principles
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Construction, installation and protection of cables and other elements of outside plant
Series M	TMN and network maintenance: international transmission systems, telephone circuits, telegraphy, facsimile and leased circuits
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks and open system communications
Series Y	**Global information infrastructure, Internet protocol aspects and Next Generation Networks**
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2004