Quality of Services

1) INTRODUCTION

The flat paradigm on which the network was built is suitable for initial network .At that time there are not many multimedia applications .Routes in IP are determined by the IP destination address and the state of routing tables in each router on the path to the destination. Thus, all IP packets follow the same route until the route changes due to congestion, topology updates, etc.; thus, there is no implicit route per service possible. So to provide better quality service for multimedia there was a need to come up with QOS.

2) TOS byte

An attempt in IPv4 to classify traffic according to a Type of Service (TOS) byte in the IP header did not succeed Internet-wide because the TOS byte was based on fair self-classification of applications with respect to other application traffic. Multimedia applications were scarce at the time, so no real efforts were made to address this problem in the early stages of the Internet and the TOS byte was never used uniformly.

There are already different thoughts on the issue of identifying the right location for providing QoS.

The possible locations are:

i) End System-Based QoS

Mechanisms exist that use extrapolation of missing data, intelligent playback buffers that can compensate for variances in interarrival times . These mechanisms have been proven to work even over long Internet routes and moderately busy lines. While this type of QoS has its merits (e.g., simplicity), it does not scale well for true multimedia support.

ii) Service-Based QoS :

In IPv6, different service classes could also be implemented by different multicast groups .In this paradigm, not even an explicit priority indication is required because it is implicitly bound to the respective multicast group.The sender will have to provide basically the same data multiple times, and the receiver needs to know which group to subscribe to in order to match the specific transmission and end-system capabilities.

iii) Class/Priority-Based QoS :

Another paradigm proposes that, to handle multimedia traffic requirements, routers need explicit information on how to deal with packets with different service requirements. To provide this information to routers, IP packets have to contain corresponding information, as well as signalling information 5concerning route changes.

iv) Resource Reservation-Based QoS:

The most complex QoS paradigm is based on the assumption that routers must have full knowledge of connections and their QoS requirements to reserve sufficient resources to guarantee processing of packets within their specified QoS limits. With proper resource reservation setup negotiations and signalling, QoS can be guaranteed end-to-end between sender and receiver(s), either in a unicast model or using multicast, receiver-driven group management mechanisms. In other words, the receiver decides on the quality versus cost ratio and signals requirements upstream toward the sender, while routers are multicast-aware and forward traffic according to current group membership.

3) <u>Main concept of the RFC 2474 :</u>

Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers

Overview of The document :

Differentiated services enhancements to the Internet protocol are meant to enable scalable service discrimination in the Internet without the need for per-flow state and signalling at every hop. The services may be either end-to-end or intra-domain; they include both those that can satisfy quantitative performance requirements (e.g., peak bandwidth) and those based on relative performance (e.g., "class" differentiation).

Services can be constructed by a combination of:

1. Setting bits in an IP header field at network boundaries.
2. Using those bits to determine how packets are forwarded by the nodes inside the network,
3. Putting conditions on the marked packets at network boundaries according to the requirements or rules of each service.

What does the DS-compliant Network Node contain ?

The requirements or rules of each service must be set by the administrator [Not discussed in this RFC]. A differentiated services-compliant network node contains a classifier which classifies packets based on the value of the DS field, along with buffer management and packet scheduling mechanisms capable of delivering the specific packet forwarding treatment as per the DS field value. Setting of the DS field and conditioning of the behaviour of marked packets need only be performed at the network boundaries.

Elements defined in this RFC :

This RFC defines the IP header field, called the DS (for differentiated services) field. In IPv4, it defines the layout of the TOS octet; in IPv6, the Traffic Class octet. In addition, a base set of packet forwarding treatments, or per-hop behaviours, is defined.

A) Basic Introduction to the Differentiated Services:

Differentiated services are meant to provide a framework and building blocks to enable deployment of scalable service discrimination in the Internet. Packet forwarding is the relatively simple task that needs to be performed on a per-packet basis as quickly as possible. Forwarding uses the packet header to find an entry in a routing table that determines the packet's output interface.

Routing sets the entries in that table and may need to reflect a range of transit and other policies as well as to keep track of route failures. Routing tables are maintained as a background process to the forwarding task.

Analogously, the differentiated services architecture contains two main components. One is the fairly well-understood behaviour in the forwarding path and the other is the more complex and still emerging background policy and allocation component that configures parameters used in the forwarding path. The forwarding path behaviours include the differential treatment an individual packet receives, as implemented by queue service disciplines and/or queue management disciplines. These per-hop behaviours are required in network nodes to deliver differentiated treatment of packets irrespective of how we construct end-to-end or intra-domain services.

Per-hop behaviours and mechanisms to select them on a per-packet basis can be deployed in network nodes today and it is this aspect

of the differentiated services architecture that is being addressed first. In addition, the forwarding path may require that some monitoring, policing, and shaping be done on the network traffic designated for "special" treatment in order to enforce requirements associated with the delivery of the special treatment.

It is possible to deploy useful differentiated services in networks by using simple policies and static configurations.

Focus of this RFC :

This RFC concentrates on the forwarding path component. In the packet forwarding path, differentiated services are realized by mapping the codepoint contained in a field in the IP packet header to a particular forwarding treatment, or per-hop behaviour (PHB), at each network node along its path. The codepoints may be chosen from a set of mandatory values defined later in this RFC , from a set of recommended values to be defined in future documents, or may have purely local meaning. PHBs are expected to be implemented by employing a range of queue service and/or queue management disciplines on a network node's output interface queue: for example weighted round-robin (WRR) queue servicing or drop-preference queue management.

Role of Traffic Conditioners:

Marking is performed by traffic conditioners at network boundaries, including the edges of the network (first-hop router or source host) and administrative boundaries. Traffic conditioners may include the primitives of marking, metering, policing and shaping. Services are realized by the use of particular packet classification and traffic conditioning mechanisms at boundaries coupled with the concatenation of per-hop behaviours along the transit path of the traffic. A goal of the differentiated services architecture is to specify these building blocks for future extensibility, both of the number and type of the building blocks and of the services built from them.

B) Terminology Used in This RFC

- **Behaviour Aggregate**: a collection of packets with the same codepoint crossing a link in a particular direction. The terms "aggregate" and "behaviour aggregate" are used interchangeably in this document.

- **Classifier:** an entity which selects packets based on the content of packet headers according to defined rules.

- **Class Selector Codepoint:** any of the eight codepoints in the range ' xxx000' (where 'x' may equal '0' or '1').

- **Class Selector Compliant PHB:** A per-hop behaviour satisfying the Class Selector PHB Requirements specified .

- **Codepoint**: A specific value of the DSCP portion of the DS field. Recommended codepoints SHOULD map to specific, standardized PHBs. Multiple codepoints MAY map to the same PHB.

- **Differentiated Services Boundary**: The edge of a DS domain, where classifiers and traffic conditioners are likely to be deployed. A differentiated services boundary can be further sub-divided into ingress and egress nodes, where the ingress/egress nodes are the downstream/upstream nodes of a boundary link in a given traffic direction. A differentiated services boundary typically is found at the ingress to the first-hop differentiated services-compliant router (or network node) that a host's packets traverse, or at the egress of the last-hop differentiated services-compliant router or network node that packets traverse before arriving at a host. This is sometimes referred to as the boundary at a leaf router. A differentiated services boundary may be co-located with a host, subject to local policy. Also DS boundary.

- **Differentiated Services-Compliant**: in compliance with the requirements specified in this document. Also DS-compliant.

- **Differentiated Services Domain**: a contiguous portion of the Internet over which a consistent set of differentiated services policies are administered in a coordinated fashion. A differentiated services domain can represent different administrative domains or autonomous systems, different trust regions, different network technologies (e.g., cell/frame), hosts and routers, etc. Also DS domain.

- **Differentiated Services Field**: the IPv4 header TOS octet or the IPv6 Traffic Class octet when interpreted in conformance with the

definition given in this document. Also DS field.

- **Mechanism:** The implementation of one or more per-hop behaviours according to a particular algorithm.

- **Microflow:** a single instance of an application-to-application flow of packets which is identified by source address, destination address, protocol id, and source port, destination port (where applicable).

- **Per-hop Behaviour (PHB):** a description of the externally observable forwarding treatment applied at a differentiated services-compliant node to a behaviour aggregate. The description of a PHB SHOULD be sufficiently detailed to allow the construction of predictable services, as documented in [ARCH].

- **Per-hop Behaviour Group:** a set of one or more PHBs that can only be meaningfully specified and implemented simultaneously, due to a common constraint applying to all PHBs in the set such as a queue servicing or queue management policy. Also PHB Group.

- **Traffic Conditioning:** control functions that can be applied to a behaviour aggregate, application flow, or other operationally useful subset of traffic, e.g., routing updates. These MAY include metering, policing, shaping, and packet marking. Traffic conditioning is used to enforce agreements between domains and to condition traffic to receive a differentiated service within a domain by marking packets with the appropriate codepoint in the DS field and by monitoring and altering the temporal characteristics of the aggregate where necessary.

- **Traffic Conditioner**: an entity that performs traffic conditioning functions and which MAY contain meters, policers, shapers, and markers. Traffic conditioners are typically deployed in DS boundary nodes (i.e., not in interior nodes of a DS domain).

- **Service:** a description of the overall treatment of (a subset of) a customer's traffic across a particular domain, across a set of interconnected DS domains, or end-to-end. Service descriptions

are covered by administrative policy and services are constructed by applying traffic conditioning to create behaviour aggregates which experience a known PHB at each node within the DS domain. Multiple services can be supported by a single per-hop behaviour used in concert with a range of traffic conditioners.

Summary so far:

To summarize, classifiers and traffic conditioners are used to select which packets are to be added to behaviour aggregates. Aggregates receive differentiated treatment in a DS domain and traffic conditioners MAY alter the temporal characteristics of the aggregate to conform to some requirements. A packet's DS field is used to designate the packet's behaviour aggregate and is subsequently used to determine which forwarding treatment the packet receives. A behaviour aggregate classifier which can select a PHB, for example a differential output queue servicing discipline, based on the codepoint in the DS field SHOULD be included in all network nodes in a DS domain.

C) Differentiated Services Field Definition :

A replacement header field, called the DS field, is defined, which is intended to supersede the existing definitions of the IPv4 TOS octet and the IPv6 Traffic Class octet .

Six bits of the DS field are used as a codepoint (DSCP) to select the PHB a packet experiences at each node. A two-bit currently unused (CU) field is reserved and its definition and interpretation are outside the scope of this document. The value of the CU bits are ignored by differentiated services-compliant nodes when determining the per-hop behavior to apply to a received packet.

The DS field structure is presented below:

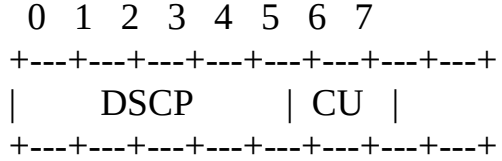

Figure (i)

DSCP: differentiated services codepoint
CU: currently unused

In a DSCP value notation 'xxxxxx' (where 'x' may equal '0' or

'1') used in this document, the left-most bit signifies bit 0 of the DS field (as shown above), and the right-most bit signifies bit 5.

Implementers should note that the DSCP field is six bits wide. DS- compliant nodes MUST select PHBs by matching against the entire 6-bit DSCP field, e.g., by treating the value of the field as a table index which is used to select a particular packet handling mechanism which has been implemented in that device. The value of the CU field MUST be ignored by PHB selection. The DSCP field is defined as an unstructured field to facilitate the definition of future per-hop behaviours.

With some exceptions noted below, the mapping of codepoints to PHBs MUST be configurable. A DS-compliant node MUST support the logical equivalent of a configurable mapping table from codepoints to PHBs. PHB specifications MUST include a recommended default codepoint, which MUST be unique for codepoints in the standard space. Implementations should support the recommended codepoint-to-PHB mappings in their default configuration. Operators may choose to use different codepoints for a PHB, either in addition to or in place of the recommended default. Note that if operators do so choose, re- marking of DS fields may be necessary at administrative boundaries even if the same PHBs are implemented on both sides of the boundary.

The exceptions to general configurability are for codepoints 'xxx000' and are also given in this RFC.

Packets received with an unrecognized codepoint SHOULD be forwarded as if they were marked for the Default behaviour, and their codepoints should not be changed. To ensure that such packets MUST NOT cause the network node to malfunction.

The structure of the DS field shown above is incompatible with the existing definition of the IPv4 TOS octet . The presumption is that DS domains protect themselves by deploying re- marking boundary nodes, as should networks using Precedence designations[Discussed in RFC 791]. Correct operational procedure SHOULD follow, which states:

"If the actual use of these precedence designations is of concern to a particular network, it is the responsibility of that network to control the access to, and use of, those precedence designations."

Validating the value of the DS field at DS boundaries is sensible in any case since an upstream node can easily set it to any arbitrary value. DS domains that are not isolated by suitably configured boundary nodes may deliver unpredictable service.

Nodes MAY rewrite the DS field as needed to provide a desired local or end-to-end service. Specifications of DS field translations at DS

boundaries are the subject of service level agreements between providers and users, [Not discussed in this RFC]. Standardized PHBs allow providers to build their services from a well-known set of packet forwarding treatments that can be expected to be present in the equipment of many vendors.

D) Historical Codepoint Definitions and PHB Requirements

The DS field will have a limited backwards compatibility with current practice, backwards compatibility is addressed in two ways. First, there are per-hop behaviours that are already in widespread use (e.g., those satisfying the IPv4 Precedence queuing requirements), and we wish to permit their continued use in DS-compliant nodes. In addition, there are some codepoints that correspond to historical use of the IP Precedence field and we reserve these codepoints to map to PHBs that meet the general requirements specified, though the specific differentiated services PHBs mapped to by those codepoints MAY have additional specifications. No attempt is made to maintain backwards compatibility with the "DTR" or TOS bits of the IPv4 TOS octet.

E) A Default PHB

A "default" PHB MUST be available in a DS-compliant node. This is the common, best-effort forwarding behaviour available in existing routers. When no other agreements are in place, it is assumed that packets belong to this aggregate. Such packets MAY be sent into a network without adhering to any particular rules and the network will deliver as many of these packets as possible and as soon as possible, subject to other resource policy constraints. A reasonable implementation of this PHB would be a queuing discipline that sends packets of this aggregate whenever the output link is not required to satisfy another PHB. A reasonable policy for constructing services would ensure that the aggregate was not "starved". This could be enforced by a mechanism in each node that reserves some minimal resources (e.g, buffers, bandwidth) for Default behaviour aggregates. This permits senders that are not differentiated services-aware to continue to use the network in the same manner as today. The RECOMMENDED codepoint for the Default PHB is the bit pattern ' 000000'; the value '000000' MUST map to a PHB that meets these specifications. The codepoint chosen for Default behaviour is compatible with existing practice . Where a codepoint is not mapped to a standardized or local use PHB, it SHOULD be mapped to the Default PHB.

A packet initially marked for the Default behaviour MAY be re-marked with another codepoint as it passes a boundary into a DS domain so that it will be forwarded using a different PHB within that domain,

possibly subject to some negotiated agreement between the peering domains.

F) Use of IP Precedence Field in Future:

We wish to maintain some form of backward compatibility with present uses of the IP Precedence Field: bits 0-2 of the IPv4 TOS octet. Routers exist that use the IP Precedence field to select different per-hop forwarding treatments in a similar way to the use proposed here for the DSCP field. Thus, a simple prototype differentiated services architecture can be quickly deployed by appropriately configuring these routers. Further, IP systems today understand the location of the IP Precedence field, and thus if these bits are used in a similar manner as DS-compliant equipment is deployed, significant failures are not likely during early deployment. In other words, strict DS-compliance need not be ubiquitous even within a single service provider's network if bits 0-2 of the DSCP field are employed in a manner similar to, or subsuming, the deployed uses of the IP Precedence field.

i) History and Evolution of IP Precedence:

The IP Precedence field is something of a forerunner of the DS field. The values that the three-bit IP Precedence Field might take were assigned to various uses, including network control traffic, routing traffic, and various levels of privilege. The least level of privilege was deemed "routine traffic". In [RFC791], the notion of Precedence was defined broadly as "An independent measure of the importance of this datagram." Not all values of the IP Precedence field were assumed to have meaning across boundaries, for instance "The Network Control precedence designation is intended to be used within a network only. The actual use and control of that designation is up to each network." [RFC791]

As networks became more complex and customer requirements grew, commercial router vendors developed ways to implement various kinds of queueing services including priority queueing, which were generally based on policies encoded in filters in the routers, which examined IP addresses, IP protocol numbers, TCP or UDP ports, and other header fields. IP Precedence was and is among the options such filters can examine.

ii) Including IP Precedence into Class Selector Codepoints :

A specification of the packet forwarding treatments selected by the IP Precedence field today would have to be quite general; probably not specific enough to build predictable services from in the differentiated services framework. To preserve partial backwards

compatibility with known current uses of the IP Precedence field without sacrificing future flexibility, we have taken the approach of describing minimum requirements on a set of PHBs that are compatible with most of the deployed forwarding treatments selected by the IP Precedence field. In addition, we give a set of codepoints that MUST map to PHBs meeting these minimum requirements. The PHBs mapped to by these codepoints MAY have a more detailed list of specifications in addition to the required ones stated here. Other codepoints MAY map to these same PHBs. We refer to this set of codepoints as the Class Selector Codepoints, and the minimum requirements for PHBs that these codepoints may map to are called the Class Selector PHB Requirements.

- **The Class Selector Codepoints**

  A specification of the packet forwarding treatments selected by The DS field values of 'xxx000|xx', or DSCP = 'xxx000' and CU subfield unspecified, are reserved as a set of Class Selector Codepoints. PHBs which are mapped to by these codepoints MUST satisfy the Class Selector PHB requirements in addition to preserving the Default PHB requirement on codepoint '000000'.

- **Requirements of The Class Selector PHB Requirements**

  We refer to a Class Selector Codepoint with a larger numerical value than another Class Selector Codepoint as having a higher relative order while a Class Selector Codepoint with a smaller numerical value than another Class Selector Codepoint is said to have a lower relative order. The set of PHBs mapped to by the eight Class Selector Codepoints MUST yield at least two independently forwarded classes of traffic, and PHBs selected by a Class Selector Codepoint SHOULD give packets a probability of timely forwarding that is not lower than that given to packets marked with a Class Selector codepoint of lower relative order, under reasonable operating conditions and traffic loads. A discarded packet is considered to be an extreme case of untimely forwarding. In addition, PHBs selected by codepoints '11x000' MUST give packets a preferential forwarding treatment by comparison to the PHB selected by codepoint '000000' to preserve the common usage of IP Precedence values '110' and '111' for routing traffic.

  Further, PHBs selected by distinct Class Selector Codepoints SHOULD be independently forwarded; that is, packets marked with different Class Selector Codepoints MAY

be re-ordered. A network node MAY enforce limits on the amount of the node's resources that can be utilized by each of these PHBs. PHB groups whose specification satisfy these requirements are referred to as Class Selector Compliant PHBs. The Class Selector PHB Requirements on codepoint '000000' are compatible with those listed for the Default PHB .

- **Using the Class Selector PHB Requirements for IP Precedence Compatibility**

  A DS-compliant network node can be deployed with a set of one or more Class Selector Compliant PHB groups. This document states that the set of codepoints 'xxx000' MUST map to such a set of PHBs. As it is also possible to map multiple codepoints to the same PHB, the vendor or the network administrator MAY configure the network node to map codepoints to PHBs irrespective of bits 3-5 of the DSCP field to yield a network that is compatible with historical IP Precedence use. Thus, for example, codepoint '011010' would map to the same PHB as codepoint '011000'.

- **Example Mechanisms for Implementing Class Selector Compliant PHB Groups**

  Class Selector Compliant PHBs can be realized by a variety of mechanisms, including strict priority queuing, weighted fair queuing (WFQ), WRR, or variants [RPS, HPFQA, DRR], or Class-Based Queuing [CBQ].

  It is important to note that these mechanisms might be available through other PHBs (standardized or not) that are available in a particular vendor's equipment. For example, future documents may standardize a Strict Priority Queuing PHB group for a set of recommended codepoints. A network administrator might configure those routers to select the Strict Priority Queuing PHBs with codepoints 'xxx000' in conformance with the requirements of this document.

  As a further example, another vendor might employ a CBQ mechanism in its routers. The CBQ mechanism could be used to implement the Strict Priority Queuing PHBs as well as a set of Class Selector Compliant PHBs with a wider range of features than would be available in a set of PHBs that did no more than meet the minimum Class Selector PHB requirements.

iii) Summary

This document defines codepoints 'xxx000' as the Class

Selector codepoints, where PHBs selected by these codepoints MUST meet the Class Selector PHB Requirements This is done to preserve a useful level of backward compatibility with current uses of the IP Precedence field in the Internet without unduly limiting future flexibility. In addition, codepoint '000000' is used as the Default PHB value for the Internet and, as such, is not configurable. The remaining seven non-zero Class Selector codepoints are configurable only to the extent that they map to PHBs that meet the requirements.

References :

- IPv6 Essentials by O'REILLY[Version 1 and Version2]
- RFC 2474
- http://www.cisco.com/en/US/docs/ios-xml/ios/qos_classn/configuration/xe-3s/ip6-qos-xe.html
- http://www.eng.uwi.tt/depts/elec/staff/rvadams/sramroop/DiffServ_Overview.htm
- http://www.6deploy.eu/tutorials/160-6deploy_IPv6_QoS_v0_2.pdf

By,
Sriharsha Karamchati