Ind Jour Math & Comp Sc. Jhs. Vol. II - Pt-I, 2013 Pg 15-25

Service Qualities of IPv6 and IPv4

Kamlesh Rana, Dr U S Pandey, Sanjay Batra
Assot. Professor, Galgotias College of Engineering, Greater Noida,
Assot. Professor, Delhi University,
ranakamles@gmail.com, uspandey1@gmail.com, us_pandey@hotmail.com,
Recceived on 3/12/2013

ABSTRACT : The term quality of service refers to the quantifiable and preferably guaranteed Quality of Network-centric Services that may be required in order to acceptably execute certain resource-sensitive applications. Such applications VoIP and Video Conferencing are sensitive to the network performance, for better service of network, the quality of service must be increased to improve the network capacity. IPv6 was developed to solve some of the problems of IPv4, such as quality of service, security and the IPv4 addresses reduction. Current Internet Protocol networks provide better traffic delivery with partial guarantees in terms of Quality of Service. There are services including voice with strong requirements for delay and delay variation, which makes it necessary to add functionality to IPv4 networks in order to gain support for such services. In this article we describe the methods with architectures that are used to provide quality of service on a network. We specify the features that use both IPv4 and IPv6 to implement quality of service. Finally, we will present the results of the comparison of two scenarios in which basic quality of service features of IPv6 and IPv4 are assessed.

Introduction

Earlier communications systems were developed as independent and isolated elements to address a particular traffic. Factors like delay, latency, bandwidth, jitter and packet loss, were not considered in this. The current trend converging IP networks is based on the development of applications that allow the exchange of all of information such as voice, data or video on the same network infrastructure.

In current trend Internet Protocol is responsible for sending and routing packets on the network with best effort to distribute packages. In this new approach a converging network must be properly designed and managed as more critical to support all types of services. The quality of service has become more popular topic because it refers to the various technologies that guarantee a certain quality over different network services.

Quality of Service

The quality of service is a collection of service requirements that must be supported by the network during data transmission. Data transmission means sending a sequence of data from source to destination for unicast or multicast. Due to quality of services it is possible to ensure proper information delivery. Quality of services gives priority to critical performance applications that share simultaneously the network resources with other non-critical applications. Quality of services in a network manages network performance and uses bandwidth more efficiently.

The term quality of service refers to the quantifiable and preferably guaranteed Quality of Network-centric Services that may be required in order to acceptably execute certain resource-sensitive applications. The term quality of service is used in the context of highly time-sensitive applications such time-sensitive applications are often classified as Real-Time applications. There are two broad categories of such RT applications:

A) Soft Real-Time applications and

B) Hard Real-Time applications

In case of Internet Protocol based networks/ internetworks, attaining quality of service is bit tricky primarily since the original Internet Protocol was not meant or designed for such a guarantee. It was primarily a Best-Effort delivery oriented model atop which the original Internet blossomed. Quality of service deals with followings:

i. Quality of service metrics

ii. Quality of service levels

iii. Architecture

i. Quality of Service Metrics

There are different metrics to measure the services provided on a network with quality of services. Most of them are defined by the working group Internet Protocol Performance Metrics. This includes such as bandwidth, amount of data transmitted per second, delay, delay

variation, cost and probability of loss among others.

A) Bandwidth

The digital bandwidth represents the amount of data that can be transmitted in a unit time. Sum of various types of delay experienced by packet during transmission is known as total delay.

B) Types of Delay

Time that take to put all the packet bits in a particular link are known as delay and these are as Time Delay, Processing Delay and Queuing Delay. The time delay that takes a bit to pass through a link on the network. The processing delay is elapsed time to process a packet in a node. The queuing delay is time-out for a packet in the queue before being transmitted.

C) Delay Variation

The delay variation measured the delay experience by the packets that come across the same route network.

D) Packet Loss

Data are transmitted on network in packets during transmission time it is necessary to be ensuring that transmitted packets are reaching at the destination so for this we measure packets transmitted on the network. So, packet loss is measured by the number of packets transmitted that are received at the destination, against the total number of packets transmitted

e) Packet Reordering

Packet reordering is considered in terms of the number of packets that are delivered to the destination in the wrong order against the total number of packets transmitted on the network.

ii. Quality of o Service Levels

Quality of service having three levels of service such as best effort, differentiated service and guaranteed service.

A) Best Effort

In this network do everything that is possible to deliver the packets to its specified destination, but there is no guarantee of happening. This model is used by applications and ftp and http.

B) Integrated Services

These are inbuilt services provides applications with a guaranteed level of service, negotiating network parameters from end to end.

C) Differential Services

A differentiated service includes a set of screening tools and collaboration mechanisms that provide certain applications or protocols depending on certain priorities over other network traffic.

D) Method of Quality of Service

To meet the requirements of service level agreements, the network must be capable to identify different types of traffic, to reserve bandwidth to prevent packet loss, to prevent and manage congestion and to prioritize traffic.

E) Congestion Management

Congestion management is a term that encompasses different types of queuing technique to handle situations where demand for bandwidth applications exceeds the total bandwidth that the network provides. Congestion managements are such as:

a) FIFO

b) Priority Queue (PQ)

c) Custom Queue (CQ)

d) Weighted Fair Queuing (WFQ)

e) Modified Deficit Round Robin (MDRR)

a) FIFO

FIFO is the simplest type of queuing and this is a simple buffer that holds outgoing packets until the transmission interface can send them. Packets are shipped out of the interface in the same order they arrive at the buffer.

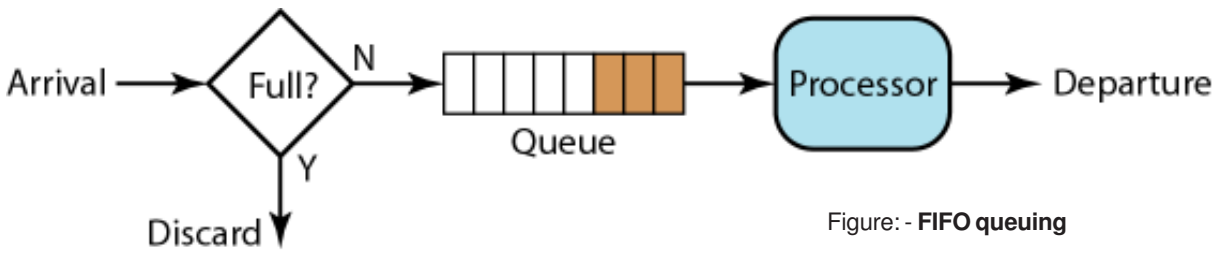

Figure: - **FIFO queuing**

b) Priority Queue (PQ)

Priority Queue is a simple approach to providing preferential action to packets. Packets arriving at the interface are separated into four lines such as low, normal, medium and high priority. The packets from the first high-priority queues are always served before others.

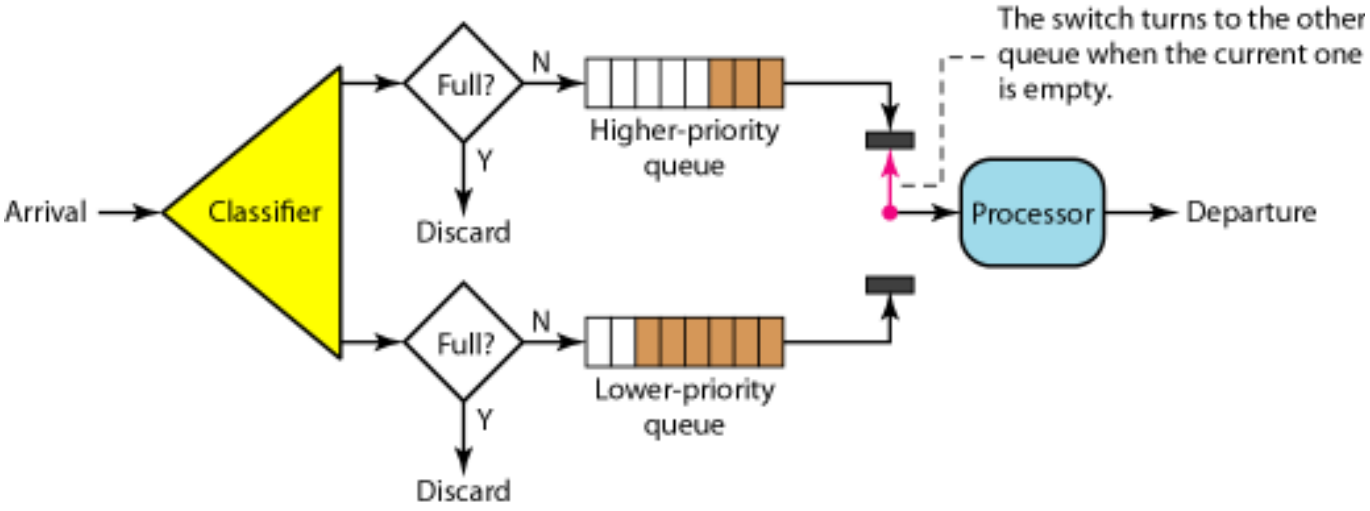

Figure: - **Priority queuing**

c) Custom Queue (CQ)

Custom Queue allows the administrator to prioritize traffics without the side effects of starvation of low priority queues, specifying the number of packets or bytes that must be addressed for each queue. Custom queue offers a more sophisticated queuing mechanism, but does not ensure a top priority as PQ does. CQ is used to provide a specific traffic guaranteed bandwidth at a point of possible congestion for traffic, ensuring a fixed portion of bandwidth and allowing other traffic to use the resources available.

d) Weighted Fair Queuing (WFQ)

Weighted fair queuing mechanism assigns weight to each packet so that traffic determines the order of packet in the queue. The weighting is done by discriminating available on TCP / IP. Weighted fair queue creates a separate queue for each type of traffic and uses a default value for queue depth.

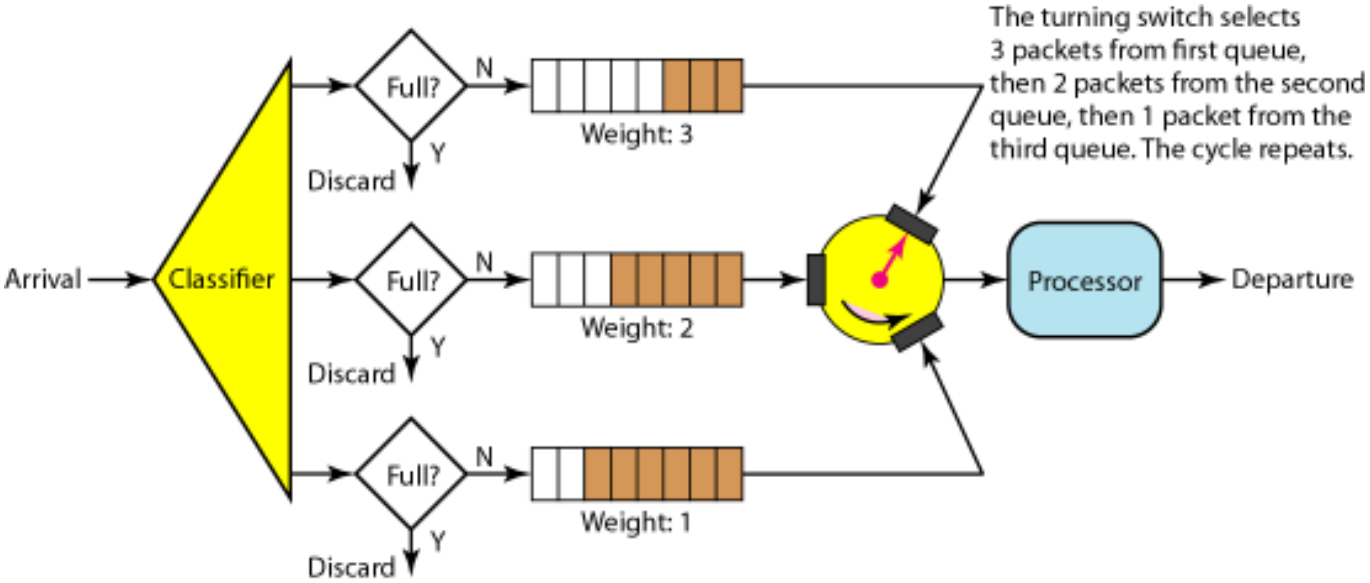

Figure: - **Weighted Fair Queuing**

e) Modified Deficit Round Robin (MDRR)

MDDR tracks the number of bytes of data that was taking out above the set value. As a result, the average amount of data served, the tail will be very close to the value set. Additionally MDDR maintains a priority queue being served preferentially.

4. Congestion Avoid Technique

Congestion avoid technique provide a set of rules for sensitive traffic. Congestion avoid technique monitor the flow of network traffic in order to anticipate and minimize its effect. There are two types of congestion avoid technique such as:

A) Random Early Detection (RED)

B) Weighted Random Early Detection (WRED)

A) Random Early Detection (RED)

RED is an active queue management scheme that provides a mechanism for congestion avoidance. Random Early Detection technique monitors the size of the queue and when queue reaches to a certain threshold, it randomly selects any TCP flows which drop packets in order to indicate the sender must reduce the rate of transmission.

RED makes two important decisions. It decides when to drop packets and what packets to drop. RED keeps track of an average queue size and drops packets when the average queue size grows beyond a defined threshold. The average size is recalculated every time a new packet arrives at the queue. RED makes packet-drop decisions based on two parameters such as Minimum Threshold (minth) and Maximum Threshold (maxth). Minimum Threshold (minth) specifies the average queue size *below which* no packets will be dropped and Maximum Threshold (maxth) specifies the average queue size *above which* all packets will be dropped

RED uses time-averaging meaning that if the queue has recently been mostly empty, RED will not react to a sudden burst as if it were a major congestion event. However, if the queues remain near full, RED will assume congestion and start dropping packets at a higher rate.

B) Weighted Random Early Detection (WRED)

It combines the capabilities of the RED algorithm with IP precedence. WRED (weighted RED) is a technique of dropping packets based on the type of traffic, where it is going, or other factors. WRED may also drop packets based on marking made to packets outside the network.

WRED avoids the globalization problems that occur when tail drop is used as the congestion avoidance mechanism on the router. Global synchronization occurs as waves of congestion crest only to be followed by troughs during which the transmission link is not fully utilized. Global synchronization of Transmission Control Protocol (TCP) hosts, for example, can occur because packets are dropped all at once. Global synchronization manifests when multiple TCP hosts reduce their transmission rates in response to packet dropping, then increase their transmission rates once again when the congestion is reduced.

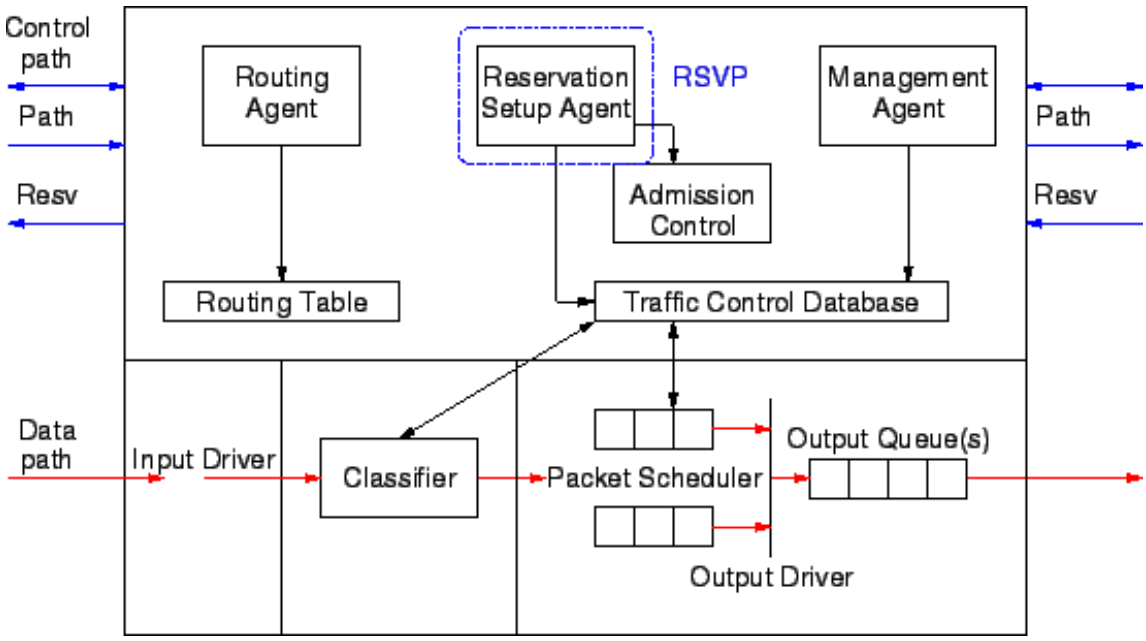

Figure:-**Integrated Services reference model with Resource Reservation Protocol (RSVP)**

iii) Architecture

There are two quality of service architectures are proposed and standardized by the IETF such as

A) Integrated Services Model (IntServ)

B) Differentiated Services (DiffServ)

Integrated Service provides quality of service through resource reservation along the data path before starting the transmission of packets.

A) Integrated Services Model (IntServ)

Integrated Services is a flow-based quality of service model designed for IP. The lack of any QoS guarantees or levels in Internet is considered as one of main limitation of more wide use of Internet.

To solve this problem an IETF Internet Integrated Services working group was formed. It defined a framework for the resource reservation and the performance guarantees. This framework is independent from protocols used for signaling and implementation details. It offers three classes of service:

a) Best Effort Service

b) Guaranteed Service

c) Controlled Load service

a) Best Effort service

In this network do everything that is possible to deliver the packets to its specified destination, but there is no guarantee of happening. This model is used by applications and ftp and http.

b) Guaranteed service

The guaranteed service is designed for applications which require certain minimum bandwidth and maximum delay. The service, as it provides firm (mathematically provable) bounds for end-to-end queuing delay, makes possible to provide service that guarantees both the delay and the bandwidth.

The traffic is considered as a fluid model: delivered queuing delays do not exceed the fluid delays by more than the specified error bounds. The maximum delay is

$$d_{\max} = \frac{b-M}{R}\frac{p-R}{p-r} + \frac{M+C_{tot}}{R+D_{tot}}, p > R \geq r$$

$$d_{\max} = \frac{M+C_{tot}}{R+D_{tot}}, r \leq p \leq R$$

where b is a token bucket depth, r is a bucket rate, p is a token bucket plus peak rate, M is a maximum datagram size, R is a bandwidth allocated to connection, C_{tot} is a end-to-end sum of rate-dependent error terms, and D_{tot} is a end-to-end sum of rate-independent, per-element error terms. When the resource reservation is being made, each node calculates its values for C and D.

Known upper bound on Delay

i) Reliable (lossless) delivery for IP packet that conform to specification,

ii) Guaranteed Bandwidth support;

c) Controlled Load service

As in the guaranteed service the service is provided for a flow confirming the same traffic specification (TSpec). The applications may assume that only very few if any packets are lost and only very few if any packets greatly exceed minimum transit delay. If a non-conforming packet is received, the network element must ensure that

1. The other controlled-load flows receive expected quality of service.

2. The excess traffic does not have an unfair impact on best-effort traffic.

3. The excess traffic is delivered best-effort basis if sufficient resources exists.

The controlled-load service provides independent the network element load the client data flow with quality of service closely approximating the quality of service the flow would receive in unloaded network. It uses capacity control to assure this.

B) Differentiated Services Model (DiffServ)

As only the Aggregate operations are realized in this case, the hallmarks of the 'Per-Flow' brand of guaranteed quality of service become largely unavailable. This leads to the state of 'no longer any explicit resource reservation' along a path in other words, no strict guarantee of quality of service for any instance of networked application is possible in this case.

One of the main problems with any resource reservation technology is burden needed for maintaining state information for each flow. In some central network nodes the number of simultaneous flows may exceed hundred thousand. If we estimate that each flow lasts for 10 seconds, there comes and goes more than 10,000 flows per second. For a reference, a large telephone exchange can handle up to million BHCA (busy hour call attempts) which equals to few hundred calls per second on average.

The number of flows makes maintaining per flow state information infeasible in core routers. The time needed to look up database entry for 5-tuple in each packet is considerable overhead compared for normal destination