Abstract—Differentiated Services (Diffserv) is the Internet architecture that uses the queuing management schemes to provision the traffic flows in the Internet backbone system. It discriminates traffic flows to a finite aggregate of classes and provides scalability solution by simplifying the complexity functions at the edge routers. In this paper, we study the end-to-end (e2e) Quality of Service (QoS) performance of File Transfer Protocol (FTP) and Constant Bit Rate (CBR) traffics transmitted over a Diffserv network. The Diffserv system applied the Token Bucket, Time Sliding Window Three Color Marker (TSW3CM) and Single Rate Three Color Marker (SRTCM) traffic provisioning policies. The e2e QoS parameters include delay, jitter, loss ratio and throughput are analyzed and compared among the policy types against the increment of traffic connections in the network system. We conclude that the FTP traffic could achieved the best overall delay performance using SRTCM policy and the best jitter performance using TSW3CM. The lowest overall loss ratio and the best throughput for FTP could be achieved using Token Bucket. Besides that, the CBR traffic has achieved the best overall delay performance using TSW3CM policy while the SRTCM policy provides the best jitter, loss ratio and throughput performances. The future works aims to design the combination of QoS aware routing, scheduling, and Diffserv queuing schemes that can adaptively maintain QoS for each type of traffic at optimum level.

Index Terms—Diffserv; Token Bucket; TSW3CM, SRTCM; QoS; NS-2.

I. INTRODUCTION

The advent of Diffserv architecture has been initiated by the Internet Engineering Task Force (IETF) a long time ago as a better solution to provide QoS guarantees in the Internet protocol (IP) networks. Compared to its predecessor like the Integrated Services (Intserv) which provides services based on per micro flow state, Diffserv outsmaps Intserv in providing better e2e QoS and preferential treatment for large heterogeneous networks system [1]. Diffserv discriminates different traffic flows which have the same commonalities to finite aggregate of classes and provides a more scalable solution for e2e QoS in IP networks by simplifying the complexity functions such as traffic classification and traffic conditioning within the edge routers [2] [3].

Previous related studies by Kaur et al in [4] have analyzed the QoS parameters for the Internet traffic using Token Bucket, Round Robin and Priority based Diffserv algorithms. The study focus on the queue buffer occupancy for each node in order to calculate the number of packet sent, number of received packets and number of lost packets. Besides that, the study in [5] has proposed a Modified an Adaptive Factor

II. SIMULATION SETUP

The simulation setup characterizes a simple ubiquitous e2e Internet system as depicted in Figure 1. The network simulations are done using NS-2.35 [8] and AWK programming tools [9]. The next subsections explain the details of network elements parameters involved in the simulations.
A. Network Configuration

The network configuration consists of 2 remote source nodes and a destination node connected via the Diffserv network as shown in Figure 1. Each remote source node is connected to the Diffserv Edge node via 10 Mbps full-duplex link with 5 ms of link delay. The Diffserv network is represented by 2 Edge nodes (i.e. ingress and egress) and a core node. The ingress (i.e. e1 node) is connected to the code via 10 Mbps full-duplex link and 5 ms link delay while the egress node (i.e. e2 node) is connected to the code via 5 Mbps full-duplex link and 5 ms of link delay. The lower link bandwidth configuration is to reflect the bottleneck of a network system. The S1 node is connected to the FTP traffic while the S2 node is connected to the CBR traffic applications. Both traffics used 1000 bytes of packet size. The CBR traffic used constant bit rate of 100 kbps per connection. The transmission rate of FTP traffic depends on the maximum congestion window size which is approximately 50 packets. Simulations are done by increasing each traffic connections from 2 until 12 for each Diffserv policy.

B. Diffserv Queues Configurations

The Diffserv used random early detection (RED) queue in edge and core routers. The RED queue consists of a single physical queue and 3 virtual queues to represent 3 per-hop-behavior (PHB) code precedences. The first level of virtual queue precedence (i.e. PHB code point 10) has the minimum and maximum buffer size thresholds of 20 and 40 packets respectively with 0.02 packet dropped probability. The second level of virtual queue (i.e. PHB codepoint 11) has the minimum and maximum buffer size thresholds of 10 and 20 packets respectively with 0.1 probability of packets drop. The lowest level of virtual queue precedence (i.e. PHB codepoint 12) has the minimum and maximum buffer size thresholds of 5 and 10 packets respectively with 0.4 probability of packets drop. The threshold and drop probability configurations are set in such a way to enforce strict transmission rate regulations within the Diffserv network.

Table 1 shows the parameter values used in the simulations for each Diffserv policy. The TSW3CM uses committed information rate (CIR) and peak information rate (PIR) with 3 drop precedences. Packet will be marked as ‘green’ if the rate is below CIR, ‘yellow’ if the rate is between CIR and PIR, and ‘red’ if above PIR values [10]. Token Bucket policy uses CIR and committed burst size (CBS) with 2 drop precedences. Similarly with TSW3CM, packets will be marked as ‘green’ if the rate is lower than CIR, ‘yellow’ if between CIR and CBS, and ‘red’ if above CBS [11]. The SRTCM policy uses CIR, CBS and excess burst size (EBS) with 3 drop precedences. Packets will be marked as ‘green’ if the rate is below CIR, ‘yellow’ if between CBS and EBS, and ‘red’ if above EBS [12].

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Source 1 (Node S1)</th>
<th>Source 2 (Node S2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIR (bps)</td>
<td>1000000</td>
<td>2000000</td>
</tr>
<tr>
<td>CBS (bytes)</td>
<td>5000</td>
<td>10000</td>
</tr>
<tr>
<td>EBS (bytes)</td>
<td>3000</td>
<td>6000</td>
</tr>
<tr>
<td>PIR (bps)</td>
<td>5000</td>
<td>10000</td>
</tr>
<tr>
<td>Traffic Type</td>
<td>FTP</td>
<td>CBR</td>
</tr>
<tr>
<td>Traffic Rate</td>
<td>(bps)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>100000</td>
<td>100000</td>
</tr>
</tbody>
</table>

The RED queue will start dropping packets randomly based on the probability of error (i.e. early drop) when the traffic’s data rate exceeds CIR or when the the burst size more than CBS and the virtual queue buffer size exceeds the previously mentioned threshold. If the traffic’s rate exceeds PIR or the burst size exceeds EBS and the virtual queue buffer size exceeds threshold, the RED queue will drop the packets excessively according to the random drop probability. The main purpose of early drop by Diffserv queues is to signal the traffic source to follow the predetermine service level requirement and to reduce the transmission rate when the network status is closer to congestion.

III. RESULTS AND DISCUSSION

The simulation results and analysis have been divided into 4 QoS categories which are the average e2e delay, jitter, packet loss ratio and throughput. The QoS parameters are calculated as the average values based on each simulation output trace file using AWK programming script and then presented in the form of graphs as shown in the next sub-sections.

A. Average End-to-End Delay

Average e2e delay is the QoS parameter often used to describe the level of service interactivity and smooth data transmission. The average e2e delays for all connections of a traffic type in a simulation, D, is calculated by summing up all of the one way connection delay, Dc, and then divided with the total number of established active connections (i.e. n parameter) during the simulation time as in (1).

$$D = \frac{\sum_{i=1}^{n} (D_{1})_{i}}{n}$$

Figure 2, 3 and 4 show the average e2e for FTP and CBR traffics over Diffserv network using Token Bucket, TSW3CM and SRTCM policies respectively. The graphs show the delay variation for each traffic type between 2 and 12 active connections. The time for each simulation round is 85 seconds.
Simulation Studies of Diffserv Policies for the Internet Traffic

Figure 2: Average end-to-end delay using Token Bucket

Figure 3: Average end-to-end delay using TSW3CM

Figure 4: Average end-to-end delay using SRTCM

Figure 5: Average accumulative delay for all Diffserv policies

B. Average End-to-End Jitter

Jitter is the e2e one way delay variation between packets transmitted from source to destination by ignoring any lost packets [13]. Jitter causes the packets to arrive at different timing and possibility in different order. Equation (2) is the general equation used to calculate jitter per connection.

\[
J(i + 1) = J(i) + (R(i + 1) - R(i)) \times (S(i + 1) - S(i))
\]  

(2)

where:

- \(S(i)\) – Time at which packet ‘I’ is transmitted from the caller.
- \(R(i)\) – Time at which packet ‘I’ is received at the receiver.

The average e2e jitter for all traffic connections in a simulation, \(J\), is then calculated by summing up all of the one way connection jitter and then divided with the total number of established active connections (i.e. \(n\) parameter) during the simulation time as shown in (3).

\[
J = \frac{\sum_{i=1}^{n}(J(x))_i}{n}
\]  

(3)

The average e2e jitters shown in Figure 6, 7 and 8 are steadily increased between 2 and 8 connections. After that the values slightly decrease for 10 and 12 active connections for TSW3CM and SRTCM. The declining trend maybe because of too many packets have been dropped in the congested links as the network moves towards saturation point and the jitter counted in the simulations ignored the lost packets. The jitter variation could severely degrade the e2e performance of delay-sensitive traffic compared to the throughput-sensitive traffic. The ITU standard recommendation for jitter depends on the application type but it is better to keep the value to the minimum for better performance. Based on Figure 6, 7 and 8, the maximum jitters are produced by FTP traffic which are approximately 9.6852 ms, 7.0357 ms and 6.1882 ms respectively.

Figure 9 shows the average accumulative jitter for all policy types. The accumulative jitter is calculated by averaging the one-way jitters in all simulations rounds using different policy types. The SRTCM policy provides the lowest delay for CBR traffic which is about 37.2197 ms and this has made it the most suitable Diffserv policy type for delay-sensitive traffic in this network configuration.
C. Average End-to-End Loss Ratio

Packet loss ratio is the ratio of total packet loss over packet sent from the traffic source. The packet loss ratio for a traffic connection in a simulation is measured as in (4).

\[
L(i) = \frac{\sum_{j=1}^{i} (P_j)}{\sum_{k=1}^{i} (P_k)}
\]

where \( \sum P_j \) is the total packets loss and \( \sum P_k \) is the total sending packets from the source during a connection session. By considering all generated connections (i.e. \( n \) parameter) in a simulation run time, the average e2e packet loss ratio is calculated as in (5).

\[
L = \frac{\sum_{i=1}^{i=n} L(i)}{n}
\]

Figure 10, 11 and 12 show the average e2e packet loss ratio using Token Bucket, TSW3CM and SRTCM policies respectively. The average e2e packet loss ratio is proportional to the increment of average generated traffic connections in the network system. The lower the loss ratio, the better would be the e2e service quality. The maximum average e2e loss ratio for FTP traffic using Token Bucket, TSW3CM and SRTCM policies are 11.51%, 12.79% and 12.55% while for CBR traffic the values are 0.38%, 11.88% and 0.08% respectively. The maximum loss ratio occurs when the maximum number of active connections generated between source and destination. In addition, the minimum average e2e loss ratio for FTP traffic using Token Bucket, TSW3CM and SRTCM policies are 0.35%, 1.41% and 1.46% respectively while for CBR traffic the values are 0%, 1.88% and 0% respectively during 2 active connections.

Figure 13 shows the average accumulative e2e loss ratio for all Diffserv policies. The values are calculated by averaging the loss ratio values in all simulation rounds for each Diffserv policy configuration. The minimum average accumulative loss ratio for FTP traffic is 6.63% using Token Bucket policy while the value for CBR traffic is 0.03% using SRTCM policy.

The packet loss occurs when the queue buffer on a link element becomes overflow as the results of network congestion. In a Diffserv network, the loss is also due to the early packet dropped by the Diffserv buffer either at the edge or core node if the transmitted rate exceeds either one of the CIR, PIR, EBS and CBS parameters. The packet loss is a critical factor for the throughput-sensitive traffic like FTP and the value must be kept at minimum. For this network...
scenario, the best Diffserv policy for throughput-sensitive traffic is Token Bucket.

\[ T = \frac{\sum_{j=1}^{n} P_j \times 8}{\sum_{j=1}^{n} t_r - t_c} \]  

(6)

where \( P_j \) is the total received packets at the receiver in bytes, \( t_r \) is the packet received time and \( t_c \) is the packet sending time from the source of a traffic connection.

Figure 11: Average end-to-end loss ratio using TSW3CM

Figure 12: Average end-to-end loss ratio using SRTCM

Figure 13: Average accumulative loss ratio for all Diffserv policies

D. Average End-to-End Throughput

The average e2e throughput summarize the previous QoS performance parameters as it measures the rate of successfully received packets at the receiver. The throughput is the critical parameter for throughput-sensitive traffic like FTP. The average e2e throughput in bit per second (bps) for all traffic connections (i.e. \( n \) parameter) in a simulation is calculated as in (6).

The average e2e throughput for traffic connections between 2 and 12. The average e2e throughput patterns are inversely proportional to the increment of average active connections. The higher the throughput the better would be the QoS. The average throughput for FTP traffic varies between 2133.43 kbps and 273.47 kbps using Token Bucket policy, 1831.14 kbps and 259.06 kbps using TSW3CM and also between 1876.28 kbps and 258.35 kbps using SRTCM policy.

Figure 14: Average end-to-end throughput using Token Bucket

Figure 15: Average end-to-end throughput using TSW3CM

Figure 16: Average end-to-end throughput using SRTCM

Figure 17 shows the average accumulative throughput for all traffic types. The average accumulative throughput is calculated by averaging the one-way throughputs in all simulation rounds using different policy types. The best average accumulative throughput that FTP could achieve is 791.46 kbps while the CBR achieved 100 kbps using SRTCM. There are 2 main factors that affecting the throughput which are the delay and packet losses. Unlike CBR, the FTP traffic uses reliable data transmission which waits for acknowledgement prior to sending the subsequent packets. This causes additional round-trip delays and retransmission if any packet loss occurs from source to destination. For this network system configuration, the Token bucket policy shown the best throughput performance for
FTP traffic while SRTCM policy shows the best performance for CBR traffic.

Figure 17: Average accumulative throughput for all Diffserv policies

IV. CONCLUSIONS AND FUTURE WORKS

This paper shows the simulation studies of e2e QoS performance for FTP and CBR traffics over a Diffserv network system. The studies compare the e2e QoS performance of both traffics using 3 different Diffserv policies which are the Token Bucket, TSW3CM and SRTCM. The network scenario is designed with the standard parameters without specific QoS improvement modification in order to make the general performance comparison.

The FTP traffic achieves the best overall delay performance using SRTCM policy which is approximately 37.5811 ms and the best jitter performance using TSW3CM which is approximately 2.5944 ms. In addition, the Token Bucket policy shows the best loss ratio and Throughput performances which are approximately 6.63% and 791.4592 kbps respectively.

Meanwhile, the CBR traffic achieved the best overall delay performance using TSW3CM which is approximately 37.2197 ms. In addition, the SRTCM policy shows the best jitter, loss ratio and throughput performances which are approximately 5.2252 ms, 0.03% and 100kbps respectively.

This paper does not dictate the best Diffserv policy for all types of applications but rather to provide a guideline for future researchers to design better Diffserv network system configurations. The future works aims at designing the combination of QoS aware routing, scheduling and Diffserv queuing schemes that can adaptively maintain each traffic type QoS requirements at optimum level.

ACKNOWLEDGMENT

The authors would like to thank the Ministry of Education Malaysia for the generous financial support under Research Acculturation Grant Scheme (R054).

REFERENCES