Volume 4, Issue 11, November 2014 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Streaming MPEG-4 Compressed Video Over DiffServ Network

Manish P Bhatt
CSE Dept. Institute of Technology,
Nirma University,
Amedabad, India

Dr. Priyanka Sharma
CSE Dept. Institute of Technology,
Nirma University,
Amedabad, India

Prof. Purvi Kansara
CSE Dept. Institute of Technology,
Nirma University,
Amedabad, India

Abstract—Efficient peer-to-peer (P2P) video streaming is a challenging task due to bursty nature of video data and bandwidth scarcity.Differentiated Services (DiffServ) is one of the leading architectures for providing quality of service in the Internet. In this paper, We used DiffServ network with WRED, that considers QoS parameter like packet loss, delay and jitter for MPEG-4 compressed video.Simulation results show that this system can greatly improve the end-to-end delivery quality of a streaming application.

Keywords—PSNR; MOS; DiffServ; WRED; Delay; Jitter

I. INTRODUCTION

Multimedia file size is large compare to other type of file because it contains text, graphics, audio and video. Multimedia file should be compress at sender side and decompress at receiver end to transfer it over network or internet. Encoders are used to compress a multimedia files. In this paper, we used MPEG-4 encoders to compress multimedia file. MPEG-4 encoders convert video signal into a series of frames. Generally, only limited change occurs between one frame and the next, so an encoder can compress the video signal significantly by transmitting only the differences. Multimedia file has three frame types. Intra, or I-frames, carry a complete video picture and are coded without reference to other frames. I-frames compress using spatial compression. A received I-frame provides the reference point for decoding a received other frames. Predictive-coded or P-frames to be coded from a preceding I-frame or P-frame using temporal compression. P-frames can provide increased compression compared to I-frames, with a P-frame typically 20 to 70 percent the size of an associated I-frame. Finally, bi-directionally predictive-coded, or B-frames, use the previous and next I-frame or P-frame as their reference points for motion compensation. B-frames provide further compression, typically 5 to 40 percent the size of an associated I-frame.

In MPEG encoding, frames are arranged into groups of pictures (GoPs) that include an I-frame and all subsequent frames leading up to the next I-frame. GoP size is defined by number of frames in GoP. typically GoP size is 15 or 12.

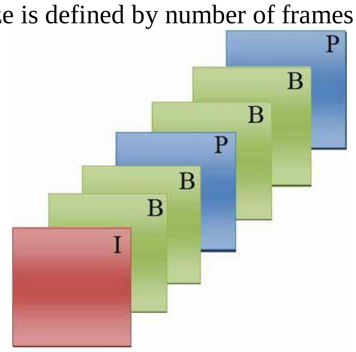

Fig. 1.Group of Picture (GoP)

This paper is organized as follows. Section II describe QoS parameter. Section III describes QoS assessment metric. Sections IV describe QoS assessment framework. Section V describes DiffServ with WRED. Section VII describes simulation. Finally section VIII concludes the paper.

II. QOS PARAMETER

Delay, jitter and packet loss are main QoS requirement for multimedia streaming. Network Delay characterizes the time difference of an IP packet between the defined network ingress point and at a defined network egress point. The IETF defines a metric for measuring one-way delay. The delays a network induces comprise four components: propagation delay along the network path, switching and queuing delays at network elements on the path, and serialization delay - that is, the time it takes to transmit the bits of the packet sequentially onto a link. Network delay can affect end-user interactivity. We must bound end-to-end network delay to bound the resulting network jitter, which affects the receiver de-jitter buffer sizing and hence impacts time. Network jitter is the variation in network delay caused by factors such as fluctuations in queuing and scheduling delays at network elements. We can generally consider jitter to be a variation of the delay. Receivers use de-jitter buffers to remove the delay variation. Video de-jitter buffer should be appropriately sized to accommodate network jitter. If the de-jitter buffer is too small to accommodate the maximum network jitter, then buffer underflows can occur - that is, the buffer will be empty when the decoder needs to process a frame, resulting in a lost packet and potential video impairment. A de-jitter buffer sized too large adds unnecessarily to the end-to-end delay, which might increase time or decrease responsiveness. Packet loss characterizes the packet drops in network. We consider a packet lost if it doesn't arrive at the specified egress point within a defined time period. The IETF defines a metric for measuring the one-way packet loss rate (PLR) Assuming the receiver de-jitter buffer is appropriately sized. Network packet loss has three primary causes: When congestion occurs, queues build up and the network drops packets. Lower-layer errors like bit errors, which might occur due to noise or attenuation in the transmission medium, can result in dropped packets. In practice, actual bit-error rates vary depending on the underlying layer-1 or layer-2 technologies used, which are different for different parts of the network. Some link-layer technologies employ reliability mechanisms, such as forward error correction (FEC), to recover from commonly occurring bit-error cases and thus reduce the effective PLR. Network element error such as link or router failures, can result in losses of network connectivity, which cause packets to be dropped until the network connectivity is restored around the failed network element.

III. QoS ASSESSMENT METRIC : PSNR AND MOS

There are basically two approaches to measure digital video quality, namely subjective quality measures and objective quality measures. Subjective quality metrics always grasp the crucial factor, the impression of the user watching the video while they are extremely costly.The human quality impression usually is given on a scale from 5 (best) to 1 (worst) as in Table 1. This scale is called Mean Opinion Score (MOS).The expensive and complex subjective tests can often not be afforded. Therefore, objective metrics have been developed to emulate the quality impression of the human visual system (HVS).

The most widespread method is the calculation of peak signal to noise ratio (PSNR) image by image. It is a derivative of the well known signal to noise ratio (SNR), which compares the signal energy to the error energy. The PSNR compares the maximum possible signal energy to the noise energy, which has shown to result in a higher correlation with the subjective quality perception than the conventional SNR.

TABLE 1

ITU-R quality and impairment scale	Scale	Quality	Impairment
	5	Excellent	Imperceptible
	4	good	Perceptible but not annoying
	3	fair	Slightly annoying
	2	poor	Annoying
	1	bad	Very annoying

Let p^{di}_{n} denote dith pixel of the nth reconstructed frame at the decoder, and p^{ei}_{n} denote eith pixel of the nth original coded frame at the encoder. The total frame error at frame n is defined as

$$e_n = \sum_{i=1}^{M} (p^{di}n - p^{ei}n) \qquad (1)$$

Where M is the number of pixels in each frame. The Mean Square Error (MSE) associated with frame error e_n is given by

$$d_n = E[(p^{di}n - p^{ei}n)^2] = \frac{1}{M}\sum_{i=1}^{M}[(p^{di}n - p^{ei}n)^2] \qquad (2)$$

Since the codec employs motion compensation and inter prediction to encode consecutive frames, distortion propagates to subsequent frames. Thus, the total distortion for a single lost frame at n is

$$D_n = \sum_{i \geq n} d_i \qquad (3)$$

The PSNR of the video signal of frame n is given by

$$PSNR_{dB}[n] = 20\log_{10}(V_{peak}/RMSE) = 20\log_{10}(V_{peak}/\sqrt{d_n}) \qquad (4)$$

Where V_{peak} is the maximum possible pixel value of the frame and RMSE is the root mean square error between received and original frames.

We consider luminance component Y of image and we represent pixel in row and column then equation 4 can be written as below

$$PSNR_{dB}[n] = 20\log_{10}\left[\frac{V_{peak}}{\sqrt{\frac{1}{N_{col}N_{row}}\sum_{i=0}^{N_{col}}\sum_{j=0}^{N_{row}}[Y_S(n,i,j) - Y_D(n,i,j)]^2}}\right] \qquad (5)$$

Where $V_{peak}=2^k-1$ and k=number of bits per pixel (luminance component). Denominator is nothing but a mean square error(MSE). PSNR measures the error between a reconstructed image and the original one. Prior to transmission, it is possible to compute a reference PSNR value sequence on the reconstruction of the encoded video as compared to the original raw video. After transmission, the PSNR is computed at the receiver for the reconstructed video of the possibly corrupted video sequence received. Conversion of PSNR to MOS shown in table 2.

TABLE 2

Possible PSNR to MOS conversion	PSNR(dB)	MOS
	> 37	5(Excellent)
	31-37	4(good)
	25-31	3(fair)
	20-25	2(poor)
	< 20	1(bad)

IV. QOS ASSESSMENT FRAMEWORK : EVALVID FRAMEWORK WITH NS2

The QoS assessment framework for video traffic combines EvalVid Framework and NS2 as shown in figure-2. These framework are designed to evaluate the quality of delivered video. The evaluation process starts from encoding the source video either in the YUV QCIF or in the YUV CIF formats. MPEG4 encoder are used to compress source video and generate compressed file. The VS program will read the compressed file and generate the traffic trace file contain frame number, frame type, frame size, number of UDP packet and sender time. The MyTrafficTrace agent extracts the frame type and the frame size of the video trace file generated from the traffic trace file, fragments the video frames into smaller segments, and sends these segments to the lower UDP layer at the appropriate time according to the user settings specified in the simulation script file. MyUDP is an extension of the UDP agent. This new agent allows users to specify the output file name of the sender trace file and it records the timestamp of each transmitted packet, the packet ID, and the packet payload size. MyUDPSink is the receiving agent for the fragmented video frame packets sent by MyUDP. This agent also records the timestamp, packet ID, and payload size of each received packet in the user specified receiver trace file. After simulation, the evaluation task begins. Based on these three trace files and the original encoded video, the ET program produces packet loss, delay, jitter report and the corrupted video file. Afterward, the corrupted video is decoded and generate reconstructed erroneous video. The total number of video frames at the receiver side, including the erroneous frames, must be the same as that of the original video at the sender side. If the codec cannot handle missing frames, the FV component is used to tackle this problem by inserting the last successfully decoded frame in the place of each lost frame as an error concealment technique. Finally, the reconstructed fixed YUV video can be compared with the original raw YUV video to evaluate the end-to-end delivered video quality. Before you begin to format your paper, first write and save the content as a separate text file. Keep your text and graphic files

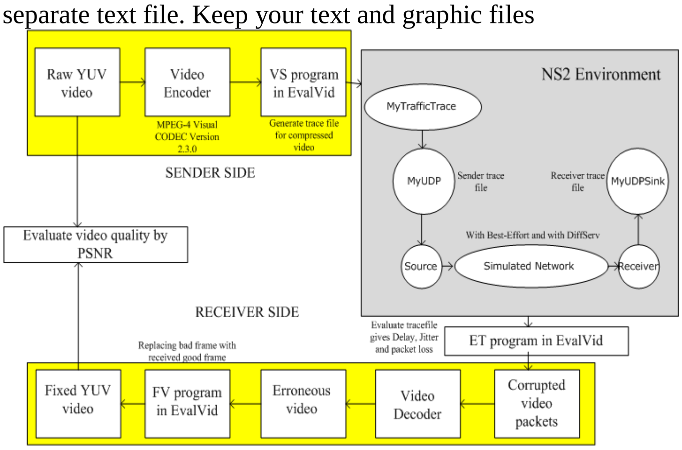

Fig. 2. QoS assessment framework

V. DIFFSERV WITH WRED

DiffServ uses codepoints(DSCP) attached to a packet's IP-header to distinguish traffic with different PHBs (Per Hop Behavior). A PHBs defines a forwarding treatment of a single packet in a router. It is defining better treatment for some class of traffic than to another. Four standard PHBs are The Default PHB, Class-Selector PHBs, Expedited Forwarding PHB, Assured Forwarding PHB.

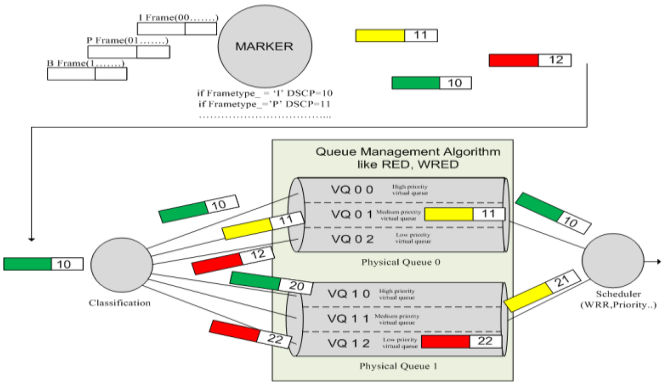

Fig. 3. DiffServ network with WRED

DiffServ manage packets inside a single queue or buffer using queue management technique like RED or WRED and scheduling between multiple queues, i.e. how bandwidth is shared between queues. In this paper, WRED queue management technique is used with DiffServ architecture. When the queue builds up and exceeds a given threshold, the WRED starts to drop packets following the specified drop probability parameters. In this paper we used mpeg4 compressed video file. In mpeg4 compressed video file, I-frame is more important compared to P-frame and B-Frame. As shown in figure-3,I-frame(green color) is pre-marked with DiffServ code point 10(AF11) that is assured forwarding PHB class 1 low drop precedence. P-frame(yellow color) is pre-marked with DiffServ code point 11(AF11 + 1). B-frame(red color) is pre-marked with DiffServ code point 12(AF12) that is assured forwarding PHB class 1 medium drop precedence. Packets are classified as per DSCP and send to flow accordingly. As shown in Fig. 3, two physical queues having three virtual queues with different priority level. I-frame packets send through high priority virtual queue while B-frame packets send through low priority virtual queue.

VI. Simulation

To evaluate the performance of the video quality over DiffServ network a number of experiments were conducted using NS2 simulator. Two standard sequences "Foreman"(QCIF, 30 fps, 300 frames) and "coastguard"(QCIF, 30 fps, 300 frames) are used for simulation. Fig. 4 presents the simple simulation topology, in which source s1 and s2 delivers a video traffic stream to destination D through routers R1, R2, R3 and R4. The bottleneck link has a capacity of 360 Kbps and is situated between router R3 and router R4. The queue limit at each router is set to 10 packets.

When video is transmitted with best-effort delivery, router R3 and R4 implement Droptail queue management technique.droptail queue management discards all the incoming packets when queue size reaches the queue limit.

When video is deliever over DiffServ network, I-, P- and B- frame packets are classified based on DSCP. I-frame packets are pre-marked with DSCP 10(AF10), P-frame packets are pre-marked with DSCP 11(AF10+1) and B-frame are pre-marked with DSCP 12(AF12).router R3 and R4 implement WRED queue management.when queue exceeds a given threshold value, the WRED starts to drop packets by specified drop probability parameters.The WRED parameters include a minimum threshold, a maximum threshold, and a maximum drop probability, i.e. minth, maxth, and Pmax.WRED parameter are specified respectively {4, 8, 0.025} for I-frame packets, {2, 4, 0.05} for P-frame packets and {1, 2, 0.1} for B-frame packets.

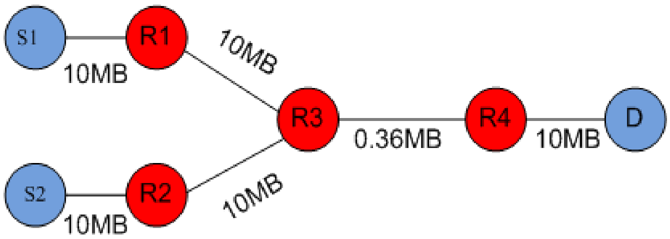

Fig. 4. Topology

Foreman video is transmitted with best-effort delievery and DiffServ network.figure-5 shows PSNR(dB) values for each frames with Best-Effort delievery and DiffServ network.we can see that using DiffServ network PSNR values are good compare to Best-Effort delievery which shows good quality frames.

Table-3 shows that with DiffSrv network we can decrease avg.Delay and avg.Jitter value compare to Best-Effort delievery and with DiffServ network we have average PSNR
value is good compare to Best-effort delievery that shows good quality frames with DiffServ network.

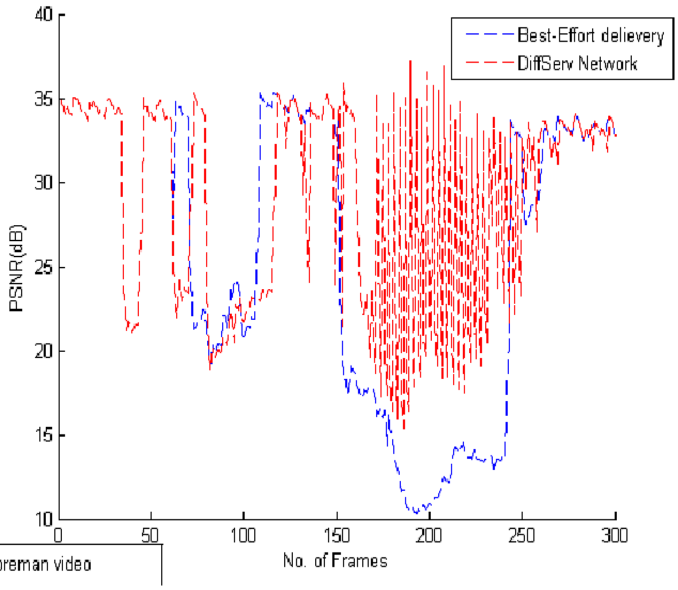

Fig. 5. PSNR value with DiffServ network and Best-effort for Foreman sequence.

TABLE 3 QoS parameter for Foreman video sequences.

video	Delay	Jitter	PSNR(dB)
Best-effort	0.195504	0.212496	25.818733
DiffServ Network	0.136452	0.158187	28.7868

0Coastguard video is transmitted with best-effort delievery and DiffServ network.figure-6 shows PSNR(dB) values for each frames with Best-Effort delivery and DiffServ network. We can see that using DiffServ network PSNR values are good compare to Best-Effort delivery which shows good quality frames.

TABLE 4 QoS parameter for Coastguard video sequences.

video	Delay	Jitter	PSNR(dB)
Best-effort	0.202501	0.22223	22.67663
DiffServ Network	0.135268	0.175043	24.3966

Table-4 shows that with DiffSrv network we can decrease avg.Delay and avg.Jitter value compare to Best-Effort delievery but average PSNR value with DiffServ network is very near to average PSNR value with Best-effort delievery for coastguard video.

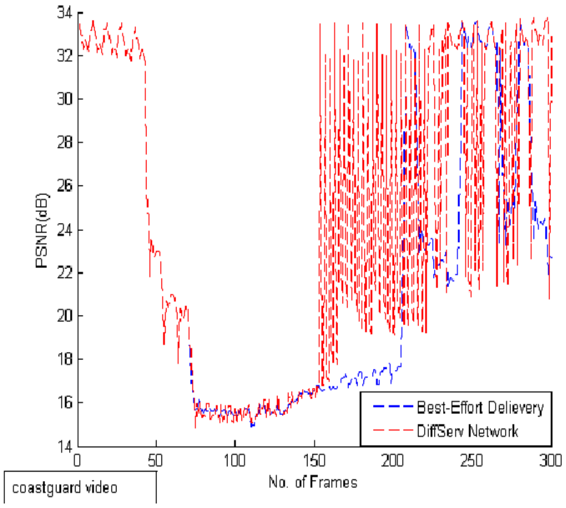

Fig. 6. PSNR value with DiffServ network and Best-effort for coastguard sequence.

VII. Conclusion

As the bursty nature of video data, it is a great challenge to guarantee video streaming QoS over internet. In this paper, we used DiffServ architecture, to classify packets and forwarding treatment as per their importance. We use WRED queue management technique to drop packets from flows as per flow priority, when queue exceeds queue limits. Simulation result shows that QoS parameter like Delay, Jitter and packet loss are optimized using DiffServ architecture compare to Best-effort delivery.

REFERENCES

[1] J. Klaue, B. Rathke, and A. Wolisz, "EvalVid A framework for video transmission and quality evaluation," in Proceedings of the International Conference on Modelling Techniques and Tools for Computer Performance Evaluation, 2003, pp. 255-272.

[2] Chih-Heng Ke, Ce-Kuen Shieh, Wen-Shyang Hwang, and Artur Ziviani,"An evaluation framework for more realistic simulations of MPEG video transmission", Journal of Information Science and Engineering,vol. 24, pp. 425–440, 2008.

[3] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "An architecture for differentiated services," RFC 2475, IETF, Dec. 1998.

[4] J. Shin, J. Kim, and C.-C. J. Kuo, "Quality-of-service mapping mechanism for packet video in differentiated services network," IEEE Trans. Multimedia, vol. 3, no. 2, pp. 219-231, Jun. 2001.

[5] C.-H. Ke, C.-K. Shieh, W.-S. Hwang, and A. Ziviani, "Two markers system for improved MPEG video delivery in a DiffServ network," IEEE Commun. Lett., vol. 9, no. 4, pp. 381-383, Apr. 2005.

[6] L. Chen, G. Liu, and F. Zhao, "An improved marking mechanism for real-time video over DiffServ networks," in Proc. PCM, Dec. 2007, pp. 510-519.

[7] Lishui Chen, and Guizhong Liu," A Delivery System for Streaming Video Over DiffServ Networks," IEEE trans. on circuits and system for video technology, vol. 20, no. 9, september 2010.