

# Measuring Interaction QoE in Internet Videoconferencing <sup>★</sup>

Prasad Calyam<sup>1</sup>, Mark Haffner<sup>1</sup>, Eylem Ekici<sup>1</sup>, Chang-Gun Lee<sup>2</sup>

<sup>1</sup> The Ohio State University, Columbus, OH 43210.

`pccalyam@oar.net`, `mhaffner@oar.net`, `ekici@ece.osu.edu`

<sup>2</sup> Seoul National University, Korea, 151-742.

`cglee@snu.ac.kr`

**Abstract.** Internet videoconferencing has emerged as a viable medium for communication and entertainment. However, its widespread use is being challenged. This is because videoconference end-users frequently experience perceptual quality impairments such as video frame freezing and voice dropouts due to changes in network conditions on the Internet. These impairments cause extra *end-user interaction effort* and correspondingly lead to unwanted *network bandwidth consumption* that affects user Quality of Experience (QoE) and Internet congestion. Hence, it is important to measure and subsequently minimize the extra end-user interaction effort in a videoconferencing system. In this paper, we describe a novel active measurement scheme that considers end-user interaction effort and the corresponding network bandwidth consumption to provide videoconferencing interaction QoE measurements. The scheme involves a “Multi-Activity Packet-Trains” (MAPTs) methodology to dynamically emulate a videoconference session’s participant interaction patterns and corresponding video activity levels that are affected by transient changes in network conditions. Also, we describe the implementation and validation of the *Vperf* tool we have developed to measure the videoconferencing interaction QoE on a network path using our proposed scheme.

## 1 Introduction

Internet videoconferencing is being used increasingly for remote meetings, distance learning, tele-medicine, etc. over the Internet. However, videoconferencing service providers are facing challenges in successfully deploying and managing large-scale videoconferencing systems on the Internet. This is because videoconference end-users frequently experience *perceptual quality impairments* such as video frame freezing and voice dropouts due to changes in network conditions i.e., “network fault events” caused by: (a) cross-traffic congestion at intermediate hops, (b) physical link fractures, and (c) last-mile bandwidth limitations.

To assist videoconferencing service providers in proactive identification and mitigation of network fault events, several metrics and active measurement tools have been developed. The metrics include network factors such as end-to-end

---

<sup>★</sup> This work has been supported in part by The Ohio Board of Regents

available bandwidth, packet delay, jitter and loss. Active measurement tools such as Ping and Iperf [1] observe the performance levels of these network factors on the Internet paths. Studies such as [2] - [6] map the different performance levels of the measured network factors to metrics that indicate severity of the perceptual quality impairments in a videoconference session. The most widely-used metric to quantify end-user Quality of Experience (QoE) in terms of perceptual quality impairment severity is the “Mean Opinion Score” (MOS).

In addition to causing perceptual quality impairments, network fault events also cause *interaction difficulties* to the videoconference participants, which the existing metrics and tools do not measure. This observation along with an explanation on the need for schemes to measure human interaction difficulties in voice and video conferences is presented in [7]. The interaction difficulties correspond to instances during a session where a ‘listening’ participant is led to interrupt a ‘talking’ participant by saying “Can you please ‘repeat’ the previous sentence?” due to a voice drop-out caused by a network fault event occurrence. In extreme cases, prolonged network fault events impair the perceptual quality and aggravate the interaction difficulties to an extent that the participants decide to ‘disconnect’ and ‘reconnect’ the videoconference session. Upon reconnection, assuming the effects of the network fault events have subsided, the participants ‘reorient’ their discussion to progress further with the remaining “agenda-items” in the videoconference session.

These ‘repeat’, ‘disconnect’, ‘reconnect’, and ‘reorient’ actions during a videoconference session are unwanted interaction patterns because they increase the end-user effort required to complete a set of agenda-items in a videoconference session. In addition, video traffic in a videoconference session has high data rates (256 Kbps - 768 Kbps) and hence the video traffic of the unwanted interaction patterns increases the Internet traffic congestion levels. The end-user interaction effort and the corresponding network bandwidth consumption together can be measured using the “agenda-bandwidth” metric, which we define as the aggregate network bandwidth consumed on both sides while completing a set of agenda-items in a videoconference session. Unwanted interaction patterns result in *unwanted agenda-bandwidth*. The ability to measure the unwanted agenda-bandwidth and subsequently minimize it using suitable traffic engineering techniques [8] can foster efficient design of large-scale videoconferencing systems.

To measure the unwanted agenda-bandwidth i.e., interaction QoE in an automated manner by mimicking the interaction behavior of participants, it is not practical to use actual videoconferencing end-points and video sequences. Hence, in this paper, we describe a *novel active measurement scheme to measure the interaction QoE of a videoconference session*. This scheme described in Section 3 (after describing a videoconferencing system in Section 2) involves a “Multi-Activity Packet-Trains” (MAPTs) methodology. Here, an interaction behavior controller illustrated in Figure 1, generates probing packet trains to dynamically emulate a videoconference session’s participant interaction patterns and corresponding video activity levels for a given session agenda input. During the session emulation, the network fault events, for example on Side-A, are detected

by analyzing online performance of the received video streams from Side-B. Such detected events affect the smooth progress of the agenda-items and cause interaction patterns with ‘repeat’, ‘disconnect’, ‘reconnect’, and ‘reorient’ actions. The interaction QoE is reported based on the session agenda progress, which is reflected in the agenda-bandwidth measurement. In Section 4, we describe the implementation of the *Vperf* tool that we have developed to measure the videoconferencing interaction QoE on a network path using our proposed scheme. In Section 5, we show *Vperf*’s measurements in a network testbed that features a wide variety of network fault events. Finally, Section 6 concludes the paper.

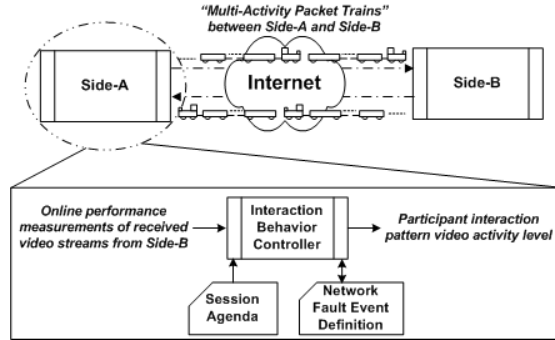


Fig. 1. MAPTs methodology

## 2 System Description

### 2.1 Video Traffic Characteristics

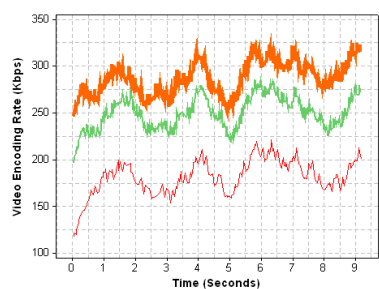
The combined voice and video traffic streams in a videoconference session are expressed in terms of sender-side encoding rate ( $b_{snd}$ ) as shown in Equation (1).

$$b_{snd} = b_{voice} + b_{video} = tps_{voice} \left( \frac{b_{codec}}{ps} \right)_{voice} + tps_{video} \left( \frac{b_{codec}}{ps} \right)_{video} \quad (1)$$

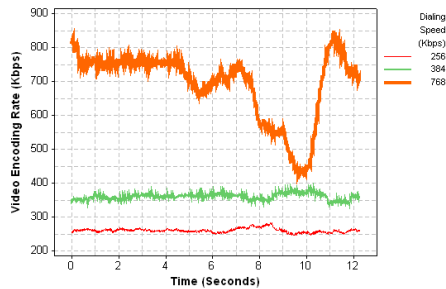
where  $tps$  corresponds to the total packet size of either voice or video packets, whose value equals a sum of the payload size ( $ps$ ), the IP/UDP/RTP header size (40 bytes) and the Ethernet header size (14 bytes);  $b_{codec}$  corresponds to the voice or video codec data rate values chosen. Commonly, G.711/G.722 voice codec and H.263 video codec are used in videoconferences. The peak encoding rate of  $\lceil b_{voice} \rceil$  is generally 64 Kbps, whereas, end-points allow end-users to choose different  $\lceil b_{video} \rceil$  settings depending on the desired video quality and the access link bandwidth at the end-user site. The end-users specify the  $\lceil b_{video} \rceil$  setting as a “dialing speed” in a videoconference. Dialing speeds of 768 Kbps

and higher are chosen by end-users with access links of T-1 or better, whereas, dialing speeds of 256 Kbps and 384 Kbps are chosen by end-users with cable-modem/DSL access links.

Owing to the sampling nature of voice codecs, voice streams contain a series of packets that are relatively small ( $tps_{voice} \leq 534$  bytes) with fixed  $ps$  characteristics. As for the video  $ps$ , they are mainly influenced by the activity-level ( $a_{lev}$ ) in the sent video sequence. This is because most video codecs use inter-frame differencing encoding, where only frames containing differences between consecutive frames are sent rather than sending every video frame. The  $a_{lev}$  refers to the temporal and spatial nature of the video sequences in a videoconference session. Broadly, video sequences can be categorized as having either low or high  $a_{lev}$ . Low  $a_{lev}$  video sequences feature slow body movements and a constant background (e.g. *Claire* video sequence). High  $a_{lev}$  video sequences feature rapid body movements and/or quick scene changes (e.g. *Foreman* video sequence). In our study, we consider the ‘listening’ end-user action to produce a low  $a_{lev}$  video sequence and the ‘talking’ end-user action to produce a high  $a_{lev}$  video sequence.



**Fig. 2.**  $b_{video}$  for *Claire* (low  $a_{lev}$ )



**Fig. 3.**  $b_{video}$  for *Foreman* (high  $a_{lev}$ )

To distinguish the low and high video  $a_{lev}$  characteristics in terms of the bandwidth consumption  $b_{video}$ , let us look at Figures 2 and 3. They show the instantaneous  $b_{video}$  values for the *Claire* and *Foreman* video sequences, respectively for the common dialing speeds. We can see that the bandwidth consumption for low  $a_{lev}$  video i.e., ‘listening’ end-user action is less than that for high  $a_{lev}$  video i.e., ‘talking’ end-user action, irrespective of the session’s dialing speed. We will use this video encoding behavior observation in our MAPTs methodology explained in Section 3.

## 2.2 Network Fault Events

The network fault events correspond to the network condition changes that affect the  $b_{snd}$  traffic and cause unwanted interaction patterns between participants in a videoconference session. The network condition changes are measured in terms

of the network factors: end-to-end network bandwidth ( $b_{net}$ ), delay ( $d_{net}$ ), jitter ( $j_{net}$ ), and loss ( $l_{net}$ ). If there is adequate  $b_{net}$  provisioned in a network path to accommodate the  $b_{snd}$  traffic, receiver-side traffic ( $b_{rcv}$ ) will be equal to  $b_{snd}$ . Otherwise,  $b_{rcv}$  is limited to  $b_{net}$  - available bandwidth at the bottleneck hop as shown in Equation (2).

$$b_{rcv} = \min(b_{snd}, \min_{i=1..hops} b_{i^{th}hop}) \quad (2)$$

Earlier studies have shown that the performance levels of network factors can be mapped to the MOS expressed in three grade ranges: [4, 5] for “Good”, [3, 4) for “Acceptable” and [1, 3) for “Poor” as shown in Table 1. Specifically, [2] and [3] suggest that for Good grade,  $b_{net}$  should be at least 20% more than the dialing speed value to accommodate the voice payload and protocol overhead;  $b_{net}$  values less than 25% of the dialing speed result in Poor Grade. The ITU-T G.114 [4] recommendation provides the levels for  $d_{net}$ . The studies in [5] and [6] provide the performance levels for  $j_{net}$  and  $l_{net}$  on the basis of empirical experiments on the Internet.

**Table 1.** Performance levels of network factors and MOS for  $[b_{video}] = 768$  Kbps

Network Factor	Good Grade	Acceptable Grade	Poor Grade
$b_{net}$	(>922] Kbps	(576-922) Kbps	[0-576) Kbps
$d_{net}$	[0-150) ms	(150-300) ms	(>300] ms
$l_{net}$	[0-0.5) %	(0.5-1.5) %	(>1.5] %
$j_{net}$	[0-20) ms	(20-50) ms	(>50] ms

In measurement studies such as [9] and [10], the transient changes of the network factors are found to occur in the form of bursts, spikes and other complex patterns and last anywhere between a few seconds to a few minutes. Considering the broad severity levels of the network fault events and the timescale within which end-point error-concealment schemes cannot ameliorate the voice and video degradation, we classify network fault events into two types: (i) Type-I, and (ii) Type-II. If the performance level of any network factor in the Good grade changes to the Acceptable grade over a 5 second period, we treat such an occurrence as a “Type-I” network fault event. Also, if the performance level of any network factor in the Good grade changes to the Poor grade over a 10 second period, we treat such an occurrence as a “Type-II” network fault event. Further, we consider the impact of a Type-I network fault event on end-user QoE to result in a ‘repeat’ action of a listening participant in a videoconference session. Along the same lines, we consider the impact of a Type-II network fault event on end-user QoE to result in the ‘disconnect’, ‘reconnect’ and ‘reorient’ actions between the participants in a videoconference session.

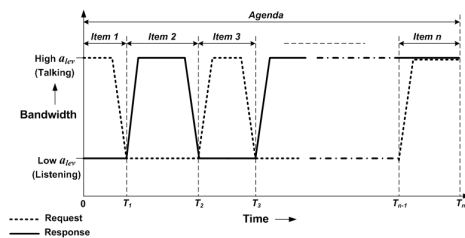
### 3 Multi-Activity Packet-Trains (MAPTs) Methodology

In this section, we explain our “Multi-Activity Packet-Trains” (MAPTs) methodology, which uses the active measurements principle where probing packet trains dynamically emulate participants’ interaction patterns and corresponding video activity levels in a videoconference session.

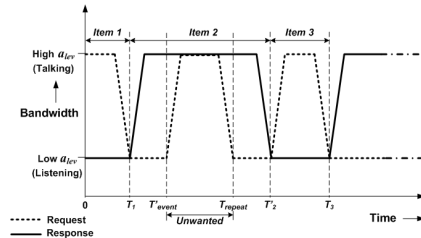
#### 3.1 Emulation of Participant Interaction Patterns

For simplicity, we assume that a videoconference session agenda involves a participant on side-A asking a series of questions to another participant on side-B. Each question or answer corresponds to a separate agenda-item. Further, if the participant on side-A is ‘talking’, we assume the participant on side-B to be ‘listening’ and vice versa. Hence, the total network bandwidth consumed by the participant on side-A asking questions, can be considered as the *request* of the videoconference session. Likewise, the total network bandwidth consumed by the participant on side-B while responding to the questions (or while satisfying the request) can be considered as the *response* of the videoconference session.

For such a videoconference session, we consider three different participant interaction patterns (PIP):  $PIP_1$ ,  $PIP_2$ , and  $PIP_3$ . The  $PIP_1$  corresponds to the participant interaction pattern when no network fault events occur during the videoconference session. Figure 4 shows the instantaneous request and response for the  $PIP_1$  interaction pattern. The videoconference session starts with the participant on side-A doing the ‘talking’ for the introduction agenda-item. Following this, the agenda-items progress with each side participant ‘talking’ alternately without any network fault event interruptions. Finally, the session ends with the both-side participants ‘talking’ during the conclusion agenda-item.



**Fig. 4.** Videoconference session request and response for  $PIP_1$

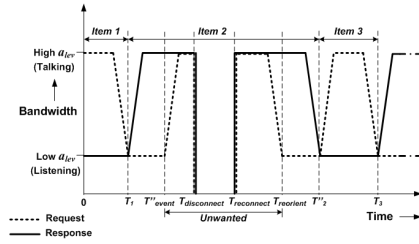


**Fig. 5.** Videoconference session request and response for  $PIP_2$

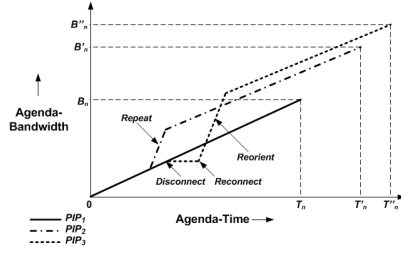
The  $PIP_2$  corresponds to the participant interaction pattern when a Type-I network fault event occurs during the videoconference session. Figure 5 shows the effects of the Type-I network fault event (occurring during agenda-item 2) on the instantaneous request and response in the videoconference session. We

can see that once the Type-I network fault event affects the ‘listening’ side-A participant at time  $T'_{event}$ , the participant begins ‘talking’ to interrupt the ‘talking’ side-B participant and requests for a repeat of the previous statements. The time between  $T'_{event}$  and  $T_{repeat}$  corresponds to the time taken for the participant on side-B to complete responding to the repeat request made by the side-A participant. The revised time to finish the item 2 in this case is  $T'_2$ .

The  $PIP_3$  corresponds to the participant interaction pattern when a Type-II network fault event occurs during the videoconference session. Figure 6 shows the effects of the Type-II network fault event (occurring during agenda-item 2) on the instantaneous request and response in the videoconference session. We can see that once the Type-II network fault event affects the ‘listening’ side-A participant at time  $T''_{event}$ , the participant begins ‘talking’ to interrupt the ‘talking’ side-B participant and requests for a session reconnection. The times  $T_{disconnect}$ ,  $T_{reconnect}$  and  $T_{reorient}$ , correspond to the ‘disconnect’, ‘reconnect’ and ‘reorient’ actions, respectively. The revised time to finish the item 2 in this case is  $T''_2$ .



**Fig. 6.** Videoconference session request and response for  $PIP_3$



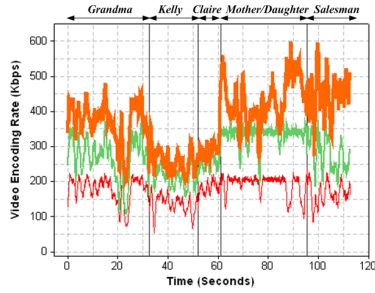
**Fig. 7.** Agenda-bandwidth for the three participant interaction patterns

Our next step is to determine the “unwanted” request and response (marked in Figures 5 and 6) and thus obtain the unwanted agenda-bandwidth in the session. Based on the agenda-bandwidth definition stated in Section 1, the agenda-bandwidth is calculated as a sum of all the instantaneous request and response bandwidth values over the session duration. The agenda-bandwidth measurements for  $PIP_1$ ,  $PIP_2$ , and  $PIP_3$  are shown in Figure 7. In the case of  $PIP_1$ , the agenda-bandwidth increases steadily and all the agenda-items consume  $B_n$  bandwidth over agenda-time  $T_n$ . We treat this agenda-bandwidth  $B_n$  over agenda-time  $T_n$  as our baseline. In the cases of  $PIP_2$  and  $PIP_3$ , the unwanted request and response increase the agenda-bandwidth to  $B'_n$  over agenda-time  $T'_n$  and  $B''_n$  over agenda-time  $T''_n$ , respectively. The goal in designing an efficient Internet videoconferencing system will be to bring the  $B'_n$  and  $B''_n$  values as close as possible to the baseline  $B_n$  using suitable traffic engineering techniques [8].

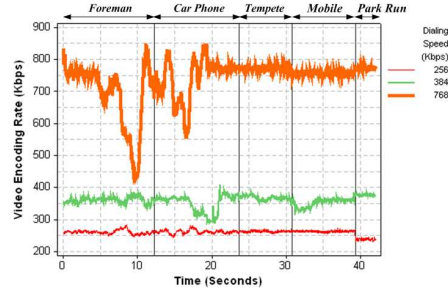
### 3.2 Emulation of Video Activity Levels

We now describe the videoconferencing traffic model to be used by the probing packet trains that emulate the low and high  $a_{lev}$  video at the different dialing speeds: 256, 384 and 768 Kbps. Since there do not exist earlier proposed videoconferencing traffic models that can be used to emulate the low and high  $a_{lev}$  video at different dialing speeds, we derive the videoconferencing traffic model parameters below using a trace-analysis based approach.

There are several video sequences available at [11] whose statistical characteristics (e.g. mean and covariance of frame quality) are widely known. Hence, researchers use them as a reference to compare the performance of their proposed techniques with other existing techniques. From these, we choose a set of 10 video sequences with 5 video sequences (*Grandma*, *Kelly*, *Claire*, *Mother/Daughter*, *Salesman*) belonging to the low  $a_{lev}$  category and 5 video sequences (*Foreman*, *Car Phone*, *Tempete*, *Mobile*, *Park Run*) belonging to the high  $a_{lev}$  category. The video sequences within a category are combined and used in a videoconference session initiated on an isolated LAN testbed with two Polycom View Station end-points that use the H.263 video codec. The  $b_{video}$  values for low and high  $a_{lev}$  traces for the common dialing speeds are shown in Figures 8 and 9, respectively. To emulate the time-series characteristics of the  $b_{video}$  values, we divide the instantaneous  $tps$  values with the corresponding instantaneous  $b_{video}$  values to obtain the instantaneous inter-packet time  $ipt$  values in the packet trains.



**Fig. 8.**  $b_{video}$  for combined low  $a_{lev}$  video sequences



**Fig. 9.**  $b_{video}$  for combined high  $a_{lev}$  video sequences

To derive the  $tps$  characteristics, we perform a statistical distribution “goodness of fit” testing. Our distribution-fit analysis suggests that the video  $tps$  distribution corresponds to the Gamma distribution given by Equation (3).

$$F(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}} \quad (3)$$

Table 2 shows the Gamma distribution shape ( $\alpha$ ) and scale ( $\beta$ ) parameters of the  $x = tps$  data in the low and high  $a_{lev}$  video traffic traces at the different dialing speeds.



To derive the trend parameters for the  $b_{video}$ , we perform time-series modeling using the classical decomposition method [12]. Our model-fit analysis suggests  $b_{video}$  as a second-order moving average (MA(2)) process given by Equation (4).

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} \quad (4)$$

where  $Z_t$  is an i.i.d. noise process with mean 0 and variance  $\sigma^2$ . Table 2 also shows the MA(2) time-series model parameters for the low and high  $a_{lev}$  video traffic at the different dialing speeds. The  $\theta_1$ ,  $\theta_2$ ,  $\mu$  and  $\sigma^2$  parameters correspond to the MA1 and MA2 co-efficients, mean and variance, respectively.

**Table 2.** Gamma distribution and MA(2) parameters for low and high  $a_{lev}$

Activity Level	Dialing Speed (Kbps)	$\alpha$	$\beta$	$\theta_1$	$\theta_2$	$\mu$	$\sigma^2$
Low	256	4.115	102.0	-1.2395	0.2395	192	200
Low	384	2.388	250.3	-1.1614	0.1614	253	400
Low	768	1.625	240.1	-1.2684	0.2684	301	500
High	256	4.321	110.0	-1.5213	0.5213	249	5
High	384	1.517	281.6	-1.4741	0.4741	349	140
High	768	1.142	446.1	-1.3024	0.3024	720	100

## 4 Vperf Implementation

In this section, we describe the implementation of the *Vperf* tool. As shown in Figure 10, the inputs to the *Vperf* tool include the session agenda and dialing speed. The session agenda consists of the *L* and *H* video  $a_{lev}$  packet train order and their lengths corresponding to the agenda-items. A simple example session agenda is shown in Figure 11. It consists of three sections that correspond to the three PIPs:  $PIP_1$ ,  $PIP_2$ , and  $PIP_3$ . Each row in the  $PIP_1$  section corresponds to a particular agenda-item. The  $PIP_2$  section row is used to specify the length of the *H* video  $a_{lev}$  packet train to emulate a ‘repeat’ action if a Type-I network fault event is detected. Similarly, the  $PIP_3$  section rows are used to specify the length and video  $a_{lev}$  of the packet trains to emulate the ‘disconnect’, ‘reconnect’ and ‘reorient’ actions if a Type-II network fault event is detected. Note that the *N* during the reconnect action corresponds to no video  $a_{lev}$  portion of the session. *Vperf* generates  $b_{video}$  from Side-A to Side-B and vice versa according to the specified session agenda and dialing speed parameters. For a particular *L* or *H* video  $a_{lev}$  specified in the session agenda, the *Vperf* tool uses the *tps* and *ipt* traffic model parameters (obtained from Section 3.2) specified in the “Traffic Model” file.

For these inputs, the outputs from *Vperf* tool are as follows: Based on the emulated traffic performance, *Vperf* continuously collects online measurements

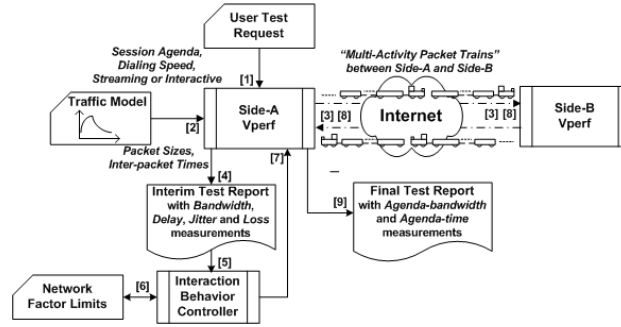


Fig. 10. Vperf tool components and their workflow

of  $b_{net}$ ,  $d_{net}$ ,  $j_{net}$ , and  $l_{net}$  network factors on a per-second basis and appends them to an “Interim Test Report”. It also produces a total average of these measurements, agenda-bandwidth and agenda-time measurements at the end of the session in the form of a “Final Test Report”. To generate these measurements for an emulated session, it uses the “Interaction Behavior Controller” component. This component processes the interim test report and detects the occurrence of Type-I and Type-II network fault events by looking up the detection rules (refer to Table 1) specified in the “Network Factor Limits” file. When a Type-I or Type-II network fault event occurrence is detected, the interaction behavior controller alters the emulation of the session agenda based on the participant interaction patterns explained in Section 3.1.

PIP-Type	Side-A Train ( $u_{Lr}$ )	Side-B Train ( $u_{Rr}$ )	Train Duration (Seconds)
PIP <sub>1</sub>	L	H	30
	H	L	10
	L	H	30
	H	L	20
	L	H	30
	H	L	20
	L	H	20
	H	L	20
PIP <sub>2</sub>	H	H	5 ← Repeat
PIP <sub>3</sub>	H	H	5 ← Disconnect
	N	N	15 ← Reconnect
	H	H	5 ← Reorient

Fig. 11. An example Vperf tool session agenda specification

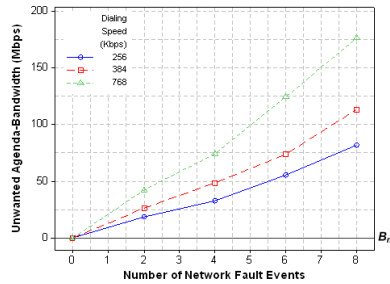


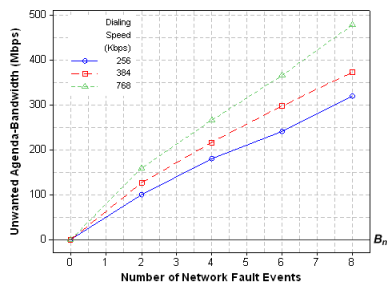
Fig. 12. Impact of increasing number of Type-I network fault events on Agenda-bandwidth

## 5 Performance Evaluation

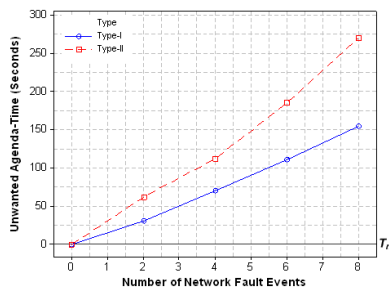
In this section, we describe the videoconference session performance measurements collected on an isolated network testbed consisting of two measurement

servers separated by the NISTnet network emulator [13]. The NISTnet was dynamically configured with different WAN profiles that featured the occurrence of varying number of Type-I and Type-II network fault events during videoconference session emulation by the Vperf tool. The session agenda input to the Vperf tool contained the packet trains video  $a_{ev}$  and train durations for the  $PIP_1$ ,  $PIP_2$  and  $PIP_3$  cases as shown in Figure 11.

Figures 12 and 13 show the unwanted agenda-bandwidth measurements from the Vperf tool for increasing number of NISTnet-generated Type-I and Type-II network fault events, respectively. Each measurement is an average of 10 emulation runs. As expected, the amount of unwanted agenda-bandwidth is zero when there is no network fault event occurrence and the unwanted agenda-bandwidth increases almost linearly with the number of network fault events occurrence. We know that the bandwidth consumed during ‘disconnect’, ‘reconnect’ and ‘reorient’ actions upon occurrence of a Type-II network fault event is higher than the bandwidth consumed in just a ‘repeat’ action upon occurrence of a Type-I network fault event. Hence, the unwanted agenda-bandwidth is greater for Type-II network fault event cases than Type-I network fault event cases, regardless of the dialing speed.



**Fig. 13.** Impact of increasing number of Type-II network fault events on Agenda-bandwidth



**Fig. 14.** Impact of increasing number of Type-I and Type-II network fault events on Agenda-time at 768 Kbps dialing speed

Similar to the unwanted agenda-bandwidth, the unwanted agenda-time also increases linearly as the number of network faults events increase. Figure 14 shows such an increase for Type-I and Type-II network fault events at the 768Kbps dialing speed. We know that the time involved in the ‘disconnect’, ‘reconnect’ and ‘reorient’ actions upon occurrence of a Type-II network fault event is higher than the time involved in just a ‘repeat’ action upon occurrence of a Type-I network fault event. Hence, we can see that the unwanted agenda-time is greater for Type-II network fault event cases than Type-I network fault event cases, regardless of the dialing speed. Similar unwanted agenda-time characteristics were observed at the 256 Kbps and 384 Kbps dialing speeds also.

## 6 Conclusion

In this paper, we presented a novel active measurement scheme used to measure the videoconferencing interaction QoE on a network path. The scheme involved a “Multi-Activity Packet-Trains” (MAPTs) methodology where probing packet trains dynamically emulated participants’ interaction patterns and corresponding video activity levels in a videoconference session that are affected by network fault events on the Internet. We detailed the characteristics of the probing packet trains with a videoconferencing traffic-model derived using a trace-analysis approach. Lastly, we described the implementation and validation of the *Vperf* tool we have developed that uses our MAPTs methodology.

Besides the basic participant interaction patterns and network fault event types considered in this paper, there are obviously several others that commonly occur in Internet videoconferencing. To formally define and classify them, detailed studies with actual end-users need to be conducted. Hence, there is a wide scope of investigation yet to be explored that can help us better understand the causes and effects of videoconference session performance failures affecting end-user interaction QoE over the Internet.

## References

1. A. Tirumala, L. Cottrell, T. Dunigan, “Measuring End-to-end Bandwidth with Iperf using Web100”, *Proc. of PAM*, 2003.
2. H. Tang, L. Duan, J. Li, “A Performance Monitoring Architecture for IP Videoconferencing”, *Proc. of IPOM*, 2004.
3. “Implementing QoS Solutions for H.323 Videoconferencing over IP”, *Cisco Systems Technical Whitepaper Document Id: 21662*, 2007.
4. *ITU-T Recommendation G.114*, “One-Way Transmission Time,” 1996.
5. P. Calyam, M. Sridharan, W. Mandrawa, P. Schopis, “Performance Measurement and Analysis of H.323 Traffic”, *Proc. of PAM*, 2004.
6. M. Claypool, J. Tanner, “The Effects of Jitter on the Perceptual Quality of Video”, *Proc. of ACM Multimedia*, 1999.
7. A. Rix, A. Bourret, M. Hollier, ”Models of Human Perception”, *BT Technology Journal - Volume 17*, 1999.
8. S. Jha, M. Hassan, ”Engineering Internet QoS”, *Artech House Publication - ISBN: 1580533418*, 2002.
9. A. Markopoulou, F. Tobagi, M. Karam, “Loss and Delay Measurements of Internet Backbones”, *Elsevier Computer Communications*, 2006.
10. L. Ciavattone, A. Morton, G. Ramachandran, “Standardized Active Measurements on a Tier 1 IP Backbone”, *IEEE Communications Magazine*, 2003.
11. Arizona State University Video Trace Library - <http://trace.eas.asu.edu>, 2007.
12. P. Brockwell, R. Davis, “Introduction to Time Series and Forecasting”, *Springer New York*, 2002.
13. NISTnet Network Emulator - <http://snad.ncsl.nist.gov/itg/nistnet>