Loss and Delay Measurements of Internet Backbones

Athina Markopoulou [a,*], Fouad Tobagi [a], Mansour Karam [b]

[a] *Electrical Engineering Dept., Stanford University, CA 94305, USA*

[b] *Avaya Inc., San Mateo, CA 94401, USA*

Abstract

As the Internet evolves into a universal network for all communication needs, it has to stand up to the high quality standards of traditional networks, such as the telephone network for voice communications. Multimedia applications are particularly sensitive to various impairments introduced by IP networks, such as packet loss, delay and delay jitter. In this paper, we study loss and delay measurements taken over the Internet and we provide a detailed characterization thereof. We focus on wide-area backbone networks, which constitute an important part of long-distance communication. Our study is based on a rich data set that provides valuable insights into the behavior of Internet backbones today, and in particular into how they affect multimedia traffic. We find that most of the problems observed seem more related to reliability, network protocols and router operation rather than to traffic load and traditional quality-of-service issues. Furthermore, the characterization and modeling of packet loss, delay and delay jitter can be used by the research community as input to various problems related to the design and evaluation of network- and application-layer mechanisms.

Key words: IP Backbone Networks, Loss, Delay and Jitter, Measurements, Characterization and Modeling

1 Introduction

As the Internet evolves into a unified network, it is important to understand its performance and capability of supporting various services at high quality. We are particularly interested in low-latency multimedia applications, such as Voice over IP (VoIP), video-conferencing and audio/video streaming. These applications are particularly demanding for two reasons: first, they have stringent requirements in terms of packet loss, end-to-end (e2e) delay and delay jitter; second, there are already high quality standards set by traditional networks, such as the telephone network for voice communications.

The first users of VoIP were eager to tolerate the bad quality because it was a free service. However, as VoIP evolves, it needs to achieve the high quality of traditional telephony. Simply stated, the problems that occur in the Internet and can affect the quality of voice and video communication are packet loss, delay and delay jitter. Loss and delay jitter can be due to congestion in the network, leading to packets getting dropped in the routers, or failure of network components leading to a reconfiguration of the network. Here the issue is how extensive are loss, delay and jitter, how bad are their effects, and whether they can be concealed at the destination.

The contribution of this paper is the collection and characterization of loss and delay measurements over a representative set of Internet backbone paths. We provide valuable insights into the behavior of Internet backbones, in particular with respect to their ability to support multimedia traffic. Furthermore, we provide a detailed characterization, and when possible modeling, of loss, delay and delay jitter, that can be used by the research community to capture the backbone network behavior.

Our study is based on a rich data set that was collected by RouteScience Technologies Inc. Probes were sent between five facilities, over a large number of different paths (43 paths belonging to 7 different Internet providers in the continental US), every 10 ms for a continuous period of 2.5 days, and accurately timestamped using GPS. The study of this data set reveals a wide range of behavior among providers: while some backbone networks exhibit excellent behavior, some other have consistent problems that severely impair the performance of multimedia traffic. Furthermore, the problems that we identify, seem more related to reliability, network protocols and router operation, rather than to traffic load and traditional Quality of Service (QoS) issues.

We focus on wide-area backbone networks, for which we have extensive data available. These are an important

* Corresponding author. Tel: +1 (650) 504-5582. Fax: +1 (650) 723-8473. Mail: 616 Harvard Ave, Apt.4, Menlo Park, CA.

Email addresses: amarko@stanfordalumni.org (Athina Markopoulou), tobagi@stanford.edu (Fouad Tobagi), mkaram@avaya.com (Mansour Karam).

part of the end-to-end path for all long distance communications, including VoIP calls that are serviced by a combination of a switched telephone network in the local area and the Internet for the long haul. Problems on the backbones will be experienced by all such calls; therefore, they need to be well understood and fixed, regardless of what takes place elsewhere in the path.

This study takes a multimedia perspective in the following sense. First in the collection phase, we sent active probes emulating voice and video traffic. Then in the characterization phase, we analyzed properties such as packet loss, delay and delay jitter which are of critical importance to interactive or streaming multimedia (as opposed to average round-trip times that would be of interest to TCP). This way, we draw conclusions about the capability of the studied networks to support multimedia applications. Furthermore, one can use our statistical characterization to capture the behavior of these networks and evaluate adaptive mechanisms at the end-systems, such as playout scheduling, multipath streaming or rate-distortion optimized streaming.

The outline of the paper is as follows. In Section 2, we review work related to network measurements and multimedia quality. In Section 3, we describe the measurement setup and collection. In Section 4 and 5, we describe the loss and delay characteristics, respectively, observed in the measurements; we provide representative examples and a detailed statistical characterization. In Section 6, we briefly discuss the effect of the observed network impairments, their possible causes and remedies. Section 7 concludes the paper.

2 Related Work

There has been an extensive amount of work on measurements and characterization of the Internet. Different studies take a different perspective depending on their specific interest (e.g. part of the Internet hierarchy under study, network protocols designed or evaluated, applications and performance metrics of interest) as well as on implementation constraints.

We are interested in the quality of multimedia communications over the public Internet. The following studies had the same objective. In [1], the delay and loss experienced by audio traffic was measured; they found that the delay variability had the form of spikes and modeled it as the result of multiplexing audio and interfering traffic into a single queue. In [11,18], audio traffic was also studied over the MBONE, and loss rates, burstiness and correlation between loss and delay were characterized. In [11], delay variability was found to have the form of spikes and playout scheduling algorithms were proposed to deal with these spikes. In [6], a large scale experiment was conducted, where low-rate MPEG-4 video was streamed to a large number of clients and cities, and statistics for the quality of the sessions were provided. Interestingly, the study of loss and delay in [12], turned out to be heavily used today in the video community, particularly for modeling network delay using Gamma distributions. Poisson flows were used in [19] to sample the network, and the constancy of delay and loss on Internet paths was studied. Finally, [4] is a recent measurement technique for inferring the state and performance of TCP-based applications based on passive measurements.

The topic of measurements from the edge of the network, is important in far too many contexts to be surveyed here exhaustively. For example, a tool for inferring ISP topologies and various metrics of interest based on measurements from the edge was developed in [16]. User-level internet-path diagnosis was provided in [10]. In general, being able to “measure the black box” is important for applications to optimize their performance.

We focus on backbone networks in the continental US, which are in general sufficiently provisioned, so they are typically believed not to introduce any impairments. Indeed, in our study we observed delay and loss patterns on those networks that seem mostly related to the network and router operation, rather than to traffic load and congestion. Similar patterns had also been observed on backbone networks in [14,15]. [5] investigated the stability and the failures of wide-area backbones due to the underlying switching system as well as due to the software and hardware components specific to the Internet’s packet-switched forwarding and routing architecture. Recent studies of the Sprint’s backbone network, [9], focused on link failures and their impact on voice traffic. They also studied the delay caused by a backbone router and identified periods during which the routers were take from serving packets [13].

A preliminary version of this paper appeared in [17]. This journal paper is significantly extended by additional materials, i.e. a complete and detailed classification and characterization of the measurements based on the work in [7]. Finally, in our previous work [8], we focused on the VoIP quality, we developed a methodology for mapping network parameters to voice subjective quality, we simulated voice calls and provided statistics on their quality. In contrast to [8], this paper focuses on the measurements themselves and on characterizing, and when possible modeling, the loss, delay and jitter observed therein. This characterization can be used by other researchers as input to problems related to the design and evaluation of network- and application-layer mechanisms.

3 Measurement Set

Our study is based on measurements provided by RouteScience Technologies Inc. Facilities have been installed in five major US cities: San Jose in California (SJC), Ashburn in Virginia (ASH), Newark in New

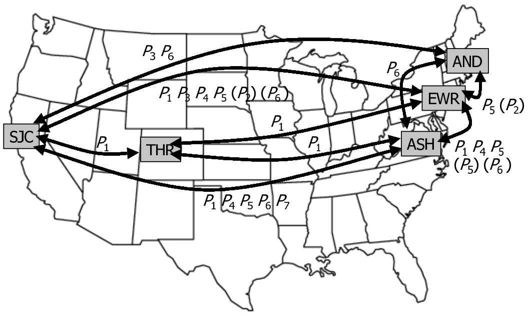


Fig. 1. Measurements collection over the backbone networks of seven major ISPs in the continental US

Jersey (EWR), Thornton in Colorado (THR) and Andover in Massachusetts (AND). These measurement facilities have been connected directly to the backbone networks of seven different Internet providers, through T1 or T3 links. We refer to the seven different providers as $P_1, P_2, ..., P_7$ for anonymity purposes. Multiple providers may connect a given pair of cities, resulting to 43 paths in total. The measurement setup is shown in Fig. 1. For example, the arrow drawn from SJC to AND with a label "P_3, P_6" means that probes were sent from SJC to AND using providers P_3 and P_6. All paths are two ways, except for those shown in parenthesis.

Probes of 50 Bytes long each were sent every 10 ms between the measurement facilities. Probes were sent from Tuesday 06/26/2001 19:22:00 until Friday 06/29/2001 00:50:00 UTC, i.e. a continuous period covering a little over two full days. "UTC" stands for Coordinated Universal Time which corresponds to Greenwich Mean Time (GMT). GPS was used to synchronize senders and receivers and the network delays were inferred by subtracting the sender's from the receiver's timestamp. The data rate of the probes (40kbps) is a very small fraction of the links used in the backbone network; therefore it could not affect the delay and loss characteristics of these networks. The size of each probe was chosen to be 50 Bytes in order to simulate a G.729 frame generated every 10 ms at 8 Kbps rate: 10B for the payload and 40B for the IP/UDP/RTP header. By taking into account the access bandwidth of the providers, we are able to compute the transmission time and infer delays for any voice packet size from the probe delays. Furthermore, the 10ms sending interval is small enough to simulate the highest rate a VoIP source might send packets at.

Sending the probes described above, we get accurate measurements of (i) the one-way delay experienced by every probe (ii) which packets are lost. This information is collected for all probes sent over 43 paths in total, belonging to 7 different providers in the continental US, over a continuous 2.5 days period. In the next sections, we describe and characterize the loss, delay and delay jitter observed in these measurements.

4 Loss Characteristics

4.1 Summary

There was only one path, namely SJC-AND for provider P3, with no loss at all during the entire measurement period. For all other paths, packet loss events with various characteristics occur. For four paths of provider P_3, loss occurred regularly for the entire measurement period, and is described in a separate section (4.4). For the remaining 38 out of the 43 paths, loss was sporadic. In general, there is no loss in the traces, except for relatively short time periods, during which, packets are lost. Therefore, it does not make sense to compute loss rates over large time periods. Indeed, no more than 0.26% of all packets are lost in any path, over the entire measurement period, but the loss rate can be much higher (10-100%) over short time periods.

We identify two types of events (i) *elementary loss events* which consist of consecutive probes getting lost (comprising one or more packets) separated by relatively long periods of time, and (ii) *complex loss events* which correspond to the occurrence of several elementary loss events concentrated over a short period of time. In the rest of this section, we give representative examples of each type. For the exhaustive list of loss events, as well as for the distributions of loss and loss-free durations on every path, the interested reader is referred to [7].

4.2 Elementary Loss Events

Elementary loss events consist of consecutive packets being lost. Their duration varies from a single packet to several consecutive packets (lasting up to tens of seconds or even a few minutes). *Single* packet loss events are a large percentage of all loss events but they contribute little to the total amount of loss. An example of 23 consecutive packets lost is shown in Fig. 2: we plot the delay incurred by probes as a function of the probe's send time; we use zero delay to indicate that a probe is lost. It is interesting to note that the pattern of 19-25 consecutive packets lost, typically preceded by high delay values, occurs frequently in providers P_2, P_3, P_5; we do not have a good explanation for the frequent occurrence of these events.

We now turn our attention to longer loss periods, which we call *outages*; these last tens of seconds up to two minutes, during which all packets are lost. The longest elementary loss event (166.18 sec) happened on the path SJC-ASH of provider P7 and is shown in Fig. 3(a). It is interesting to note that this long loss period accompanied a change in the fixed part of the delay. Also, the reverse path (from ASH to SJC) of the same provider incurred a similar loss pattern at the exact same time. The event was repeated the following day at 3:20 with loss 12.639 sec on the path SJC-ASH. Such long loss pe-

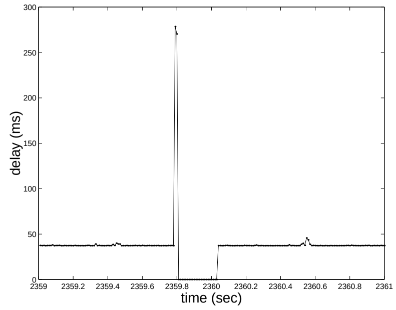
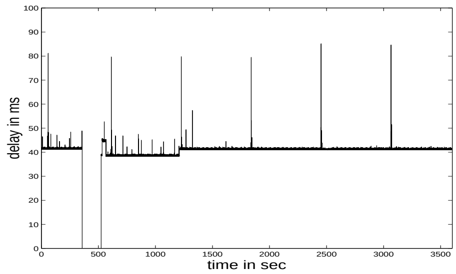
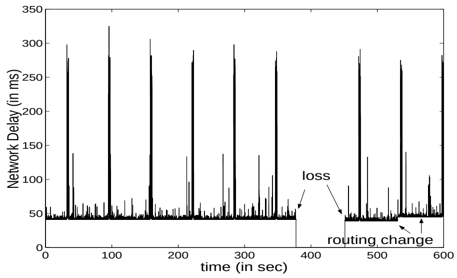


Fig. 2. Example of elementary loss (23 packets(on EWR-P2-SJC, at Thu 13:00.



(a) A 166.18sec outage, on SJC-P7-ASH, on Wed 4:00-5:00



(b) A 77.3sec outage on EWR-P4-SJC, on Wed 21:10-21:20

Fig. 3. Example Outages correlated with changes in the fixed part of the delay.

riods occurred on 6 out of 7 providers at least 1-2 times per day. For two of these providers, these outages were correlated with a change in the fixed part of the delay. The change in delay was in the order of 1-2 milliseconds, which by itself is not significant, but it indicates a reconfiguration (e.g. a routing change) that may be responsible for the long loss duration. For provider P_4, outages accompanying changes in the fixed part of the delay, was a recurrent event; an example is shown in Fig. 3(b).

4.3 Complex Loss Events

Complex loss events consist of multiple elementary loss events (single packets or longer durations) over a relatively short period of time (up to 50 seconds), during which the loss rate is 10-80%. They happen mainly on providers P_2, P_5 and P_6.

As a concrete example, we consider the path ASH-SJC of provider P6 and a single complex event that lasted 15

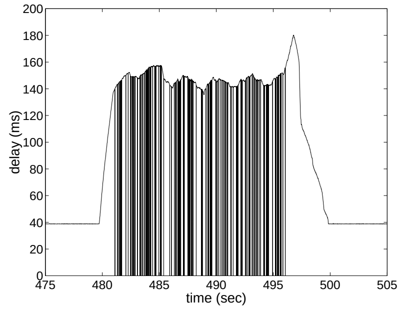


Fig. 4. Example of complex loss event consisting of single packets lost. Path ASH-P_6-SJC, Wed 06/27/01 at 3:20 (UTC). 141 packets were lost during 15 seconds: 131 single packets and 5 times two consecutive packets.

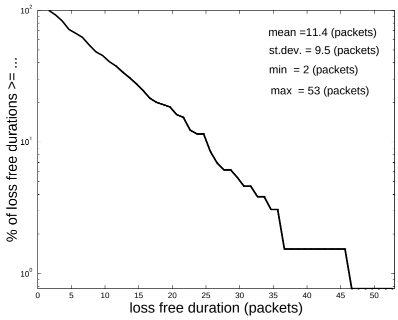


Fig. 5. Loss-free durations (i.e. times between two losses) for the complex loss event on ASH-P_6-P1, on Wed at 3:30.

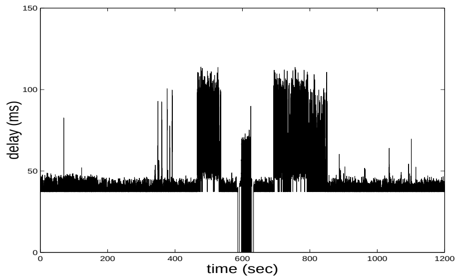
sec during which single packets were lost at a loss rate of 9.4%, see Fig. 4. The loss-free intervals are roughly exponentially distributed, as shown in Fig. 5. Indeed, the Complementary Cumulative Distribution Function (CCDF) of the exponential distribution would be a straight line in a x-logy plot. We also found that the autocorrelation function for the loss-free durations decreases fast. Similar exponential distribution of loss-free durations was also observed in the traces with regular loss (Section 4.4) as well as in other complex events.

Another event of similar type happened on EWR-P_2-SJC, on Wed at 3:30, and is shown in Fig. 6. Single packets are lost during a period of 30 seconds, (in between two longer elementary loss events lasting 502 and 512 packets each). The entire complex loss event lasts for 50 seconds, and the loss rate during that period is 24.6%; out of the 205 loss durations, 194 are single packets lost and 11 consist of 2 consecutive packets. The time between two successive losses is also exponentially distributed; we omit the statistics for lack of space.

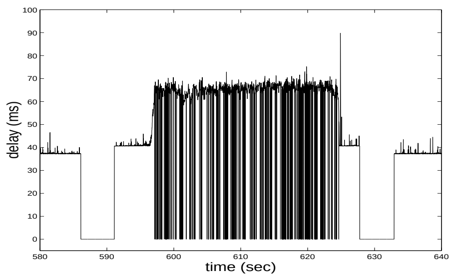
In Fig. 7, we show an example of a complex loss event consisting of longer loss durations. The event consists of longer loss durations (ranging from 10 to 143 packets, with an mean and standard deviation of 103 and 41 packets respectively). The loss-free durations range from 12 to 1891 packets, with a mean and standard deviation of 118 and 259 packets respectively. The whole event has a total duration of 30 seconds and has a packet loss rate of 41.4%. Because there are only 12 loss durations in this event, we provide the exact values for the

Table 1
Complex loss event on Thu at 20:10, on both paths of P_2. For each elementary loss event, we provide its duration and the loss-free duration until the next elementary event.

Loss duration number	1	2	3	4	5	6	7	8	9	10	11	12
	path EWR-P_2-SJC, event starting at packet sequence=26356											
loss duration (in packets)	79	143	142	10	78	142	142	142	117	90	78	79
distance from next (packets)	12	13	13	129	12	13	13	38	91	891	77	-
	path EWR-P_2-SJC, event starting at sequence=26365											
loss duration (in packets)	79	143	141	11	78	143	142	141	118	90	78	79
distance from next (packets)	12	13	12	130	11	13	13	38	89	889	78	-

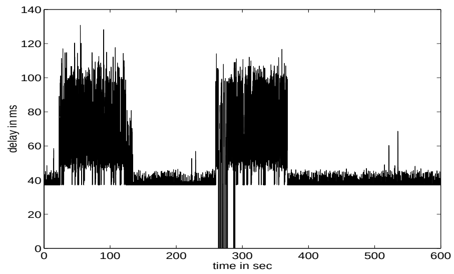


(a) 20 minutes

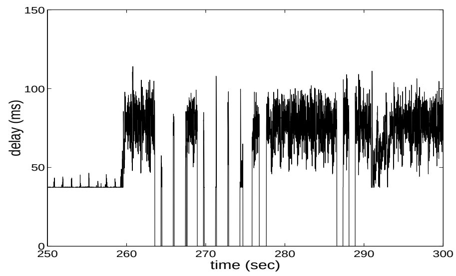


(b) Zooming in 1 minute

Fig. 6. Example of complex loss event. Path EWR-P_2-SJC, Wed 06/27/01 at 3:30-3:50 (UTC)



(a) 10 minutes



(b) Zooming in 50 sec

Fig. 7. Example of complex loss event consisting of longer loss periods. EWR-P_2-SJC, Thu 20:10.

loss durations and distances, in Table 1. Interestingly, the exact same loss pattern (same loss and loss-free durations) happened on the second path of this provider at the exact same time. A second identical event happened on both paths on Wed at 6:20. Also, the exact same event happens at the same time on the second path of provider P2. Many more occurrences of loss events happened simultaneously on many paths. The interested reader is referred to [7] for the exhaustive list.

The synchronization of loss events on many different paths indicates that these paths share a network element. Congestion on a shared link, failures or the idiosyncratic behavior of a shared router can affect all paths. The repetition of loss events with almost identical characteristics at different times on the same path, could be due to the operation of a router on this path, a maintenance or network control procedure.

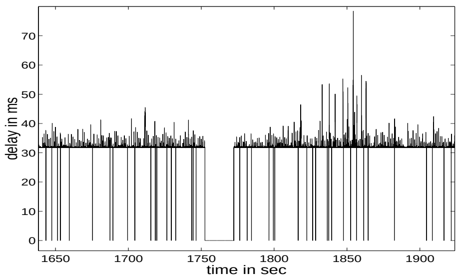
4.4 Regular Loss of Single Packets

Unlike the sporadic loss events discussed so far, four paths of provider, P_3, experience regular loss for the entire measurement period. Single packets are lost, separated by loss-free intervals; the later are exponentially distributed with an average of 5 sec.

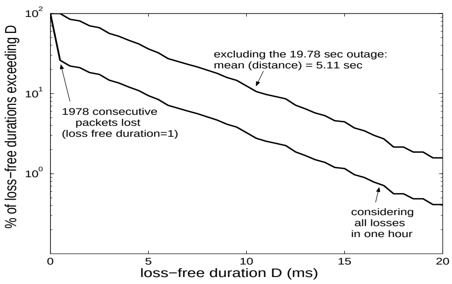
Let us look at an example path, EWR-P_3-SJC, in detail. Fig. 8 shows the CCDF of the loss-free durations, for an one-hour period. Loss during that hour, consists of single packets lost, and an outage (1978 consecutive packets lost). The intervals between single losses follow a roughly exponential distribution with mean 5.11sec; the autocorrelation function decreased fast from the first samples; the loss rate is low (0.2%). We also looked at the same path, at different times and also for longer periods. We observed the same behavior: single packets are lost, and the loss-free durations are exponentially distributed with the same mean.

5 Delay and Delay Jitter Characteristics

To aid in the analysis of delay for such a large set of measurement data, we begin by examining the statistics of delays incurred by probes over 10 minute intervals. We record for each such interval the minimum and max-



(a) Part of the one-hour period



(b) CCDF of loss-free durations

Fig. 8. One-hour period EWR-P_3-SJC, Thu 10:00-11:00.

imum delays, and various delay percentiles (primarily the 50^{th} and 99^{th} percentiles). We then plot these for all 10 minute intervals for a 24 hour period. In Fig. 9, we show such a plot for four different paths. As in the previous section, we are going to present representative delay patterns, and refer the interested reader to [7] for the exhaustive characterization.

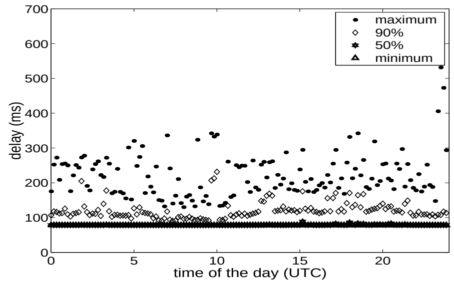
5.1 *Fixed Part of the Delay*

The minimum delay corresponds to the fixed part of the delay, which is low on the backbones under study. This is expected, as the fixed part of the delay is due to propagation and transmission delay (which is negligible on high speed backbone links; e.g. a 50 Bytes probe takes 0.266 ms on a T1 and 0.009 ms on a T3 access link). Overall, fixed delay is below 12 ms for communication on the same coast and in the range of 32-47 ms for coast-to-coast There are a few paths for which the fixed delay was as high as 78 ms, indicating that the shortest route was not followed.

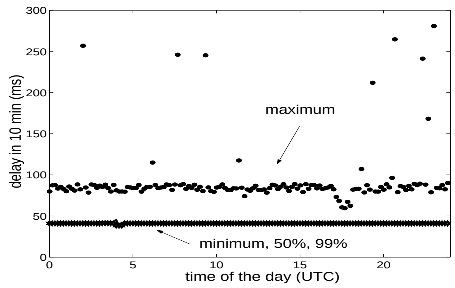
In general, the fixed part of the delay remains constant. However, we occasionally observed changes in the fixed part of the delay, in the order of a few (1-3) ms, which by itself is negligible. More frequently than not, these changes are accompanied by outages (as discussed in Section 4.2). Such changes may be due to routing changes and the outages may be due to the time it takes for network reconfiguration. However, changes in the fixed part of the delay are not always correlated with outages.

5.2 *Delay Variability*

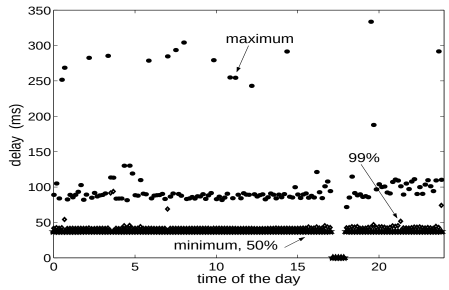
The maximum delay and delay percentiles are important for identifying intervals during which probes have



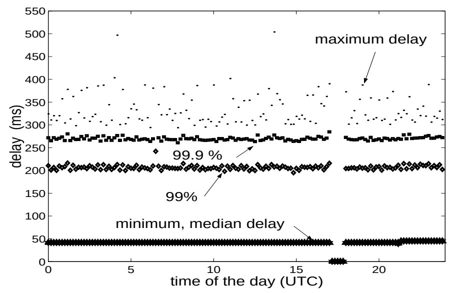
(a) THU-P1-ASH on Wed 06/27/01. A path with high delay and high delay variability



(b) SJC-P7-ASH on Wed 06/27/01. A path with low delay variability.



(c) EWR-P2-SJC on Thu 06/28/01. A path with a high and a low delay pattern.



(d) EWR-P4-SJC on Wed 06/27/01. A path with a periodic delay pattern.

Fig. 9. Delay percentiles per 10 minutes intervals for a 24 hours period and four different paths

experienced large delay. If in one 10-minute interval we observe a high maximum accompanied by increased values of the percentiles, then the interval is of interest for further study. The delay statistics exhibited in Fig. 9 are also useful to give an indication of the effect of time of day on measured delay. It also aids us in comparing paths; e.g., in Fig.9, we see that the path THU-P_1-ASH is a path that exhibits high peaks as well as high percentiles most of the day, while at the other extreme the path SJC-P_7-ASH is a path that exhibits rather low delays. The path SJC-P_2-ASH is a path that is usually good (similar to P_7) for most of the day, but incurs higher delays over a certain period of the day. The path EWR-P_4-SJC has a periodic pattern that we will discuss in detail later in this section.

We are primarily interested in analyzing the delay variations in short time scale, also called delay jitter, identifying the various possible jitter patterns and characterizing them. This requires that we plot the delay of individual probes versus their respective send times. An example is shown in Fig. 10. The delay variations that we see show that the delay is constantly varying within a certain relatively small range above the minimum. There are frequent visits to the minimum, indicating that the path are lightly loaded. This type of delay variation prevails and corresponds to what we call the normal pattern.

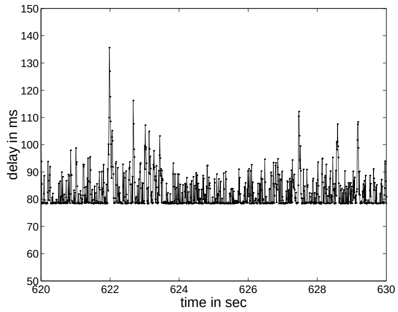
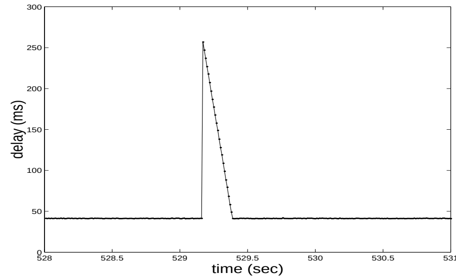


Fig. 10. Delay of individual probes on path THU-P1-ASH, on Wed 06/27/01 at 2:10

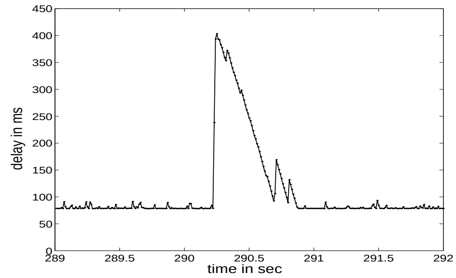
Most of the time and for most paths, the delay variability was within a few milliseconds of the fixed part. This is expected as backbone networks are usually over-provisioned with enough bandwidth to have empty queues most of the time. The lowest jitter is incurred by providers P6 and P7, for which the 99^{th} jitter percentile is from 0.1 to 0.7 ms. However, there are higher delay variations that occur mostly in the form of spikes (as opposed to a slow changing component).

By *spike* we refer to a number of packets that have significantly higher delays than the rest and they follow roughly the triangular shape shown in Fig. 11(a). There is a sudden sizable jump in delay for a probe, followed by a succession of probes delays decreasing by 10 ms each. Note that since probes are sent deterministically one every 10 ms, the delays of probes succeeding the peak follow a line with a slope of -1; this indicates that packets arrive bunched up at the receiver.

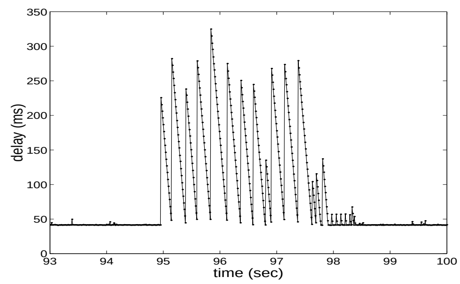
The simplest spike is the one with the perfectly triangular shape, shown in Fig. 11(a): a sudden sizable increase in delay, followed by a 45 degrees slope linear decrease. The only parameter characterizing such a spike is the magnitude of the jump, or equivalently the peak delay. The width of the spike is almost equal to the jump up to the peak delay. The spike shown in Fig. 11(b) is not as simple: there is some jitter in the decreasing slope and there are several smaller peaks that follow the first and tallest peak. In this case, the entire event may be characterized by the magnitude of the first (highest) peak, the width of the spike and the height of the smaller peaks. There are yet other situations that differ from the above description. An example is shown in Fig. 11(c): it consists of a rapid succession of spikes of similar heights lasting over three seconds. Another example is shown in Fig. 11(d): following the sudden jump in delay, a number of probes incur roughly the same delay as the peak, before the linear decrease in



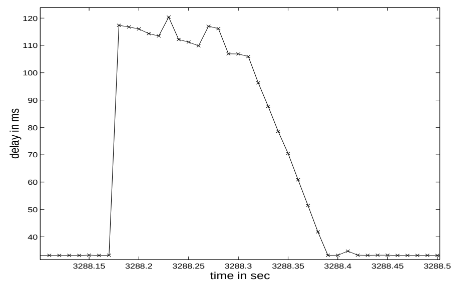
(a) "Perfect" spike, SJC-P7-ASH, Wed 2:00 (UTC)



(b) High spike, THR-P1-ASH, Wed 0:00.



(c) Clustered spikes, EWR-P4-SJC, Wed 21:00.



(d) Spike (an exception to the triangular shape), SJC-P5-EWR, Wed 17:00.

Fig. 11. Examples of Delay Spikes

delay is observed. This is an exception to the triangular spike shape, which holds in the large majority of spikes in the traces.

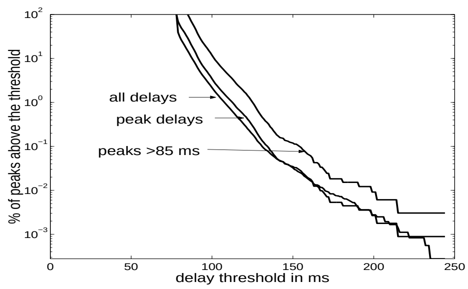
The characteristics of spikes and the specific pattern vary from path to path and over time. We illustrate this fact by examining some example paths: THR-P_1ASH, SJC-P_7-ASH and EWR-P_4-SJC. We are guided by their delay statistics for 10 minutes intervals, shown in Fig. 9 above, to select periods of time which are worth studying in greater detail. We will see that lower delays follow random patterns (consisting of spikes with random peaks at random distances) while higher delays follow periodic patterns. We finally provide a discussion on the low-frequency delay components.

5.2.1 A Path with High Delay Variability

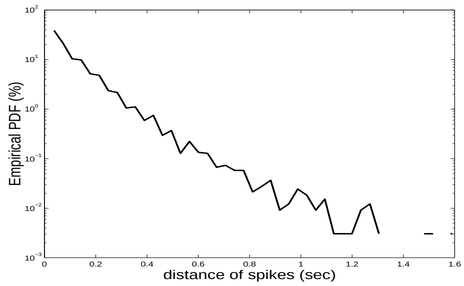
Let us first consider an example path (namely THR-P_1-ASH) with high delay and delay variability. Most of the time, delay jitter has the form of spikes of various heights spaced at various intervals from each other. Although the characteristics of the spikes vary during the day, when we zoomed in different parts of the day, we found that delay follows one of three distinct patterns. The first pattern is what we call *random delay pattern*; it holds for most of the day, when delays are relatively low. The second (*very high peaks*) and the third (*block pattern*) pattern happen when delay is high, are associated with an increase in delay percentiles in Fig. 9(a) and have some structure. We now discuss the three patterns in detail.

Random Delay Pattern. Most of the time, delay is low (roughly below 150 ms) and follows a random pattern, consisting of spikes with random peaks that happen at random intervals, as in Fig. 10. Fig. 12(a) shows the CCDF for all probe delays and for the peak delays in particular. Notice that the distribution of all probe delays is very close to the distribution of the peak delays, because of the triangular shape of the spikes. The shape of this CCDF is almost a straight line, which indicates that the exponential distribution is a good fit. We consider peak delays of a considerable size to be those above 85 ms and we observe that their distribution has also the exponential shape. The period of time separating these spikes (above 85 ms) also follows a roughly exponential distribution, as shown in Figure 12(b). The same observations hold for most of the day, when delays are small.

Very High Delays. This pattern happens when maximum delay reaches the highest values observed (e.g. 400-700 ms in Fig. 9(a) and (b), during the periods 0:00-1:00, 6:00-10:00 and 20:00-21:00, 23:00-00:00). An example of such an hour and a zoom in 50sec, is shown in Fig. 13(a) and(b) respectively; we see that these high peaks occur every 10-20 ms. Fig. 11(b) shows the details of one of these spikes: a high peak is followed by many smaller ones. When we zoomed in the remaining high

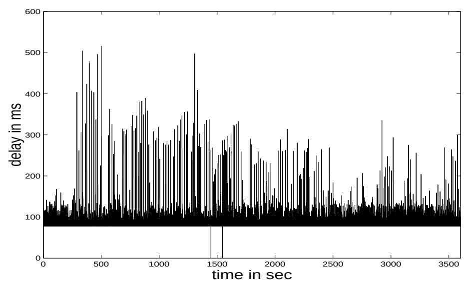


(a) Peak delays

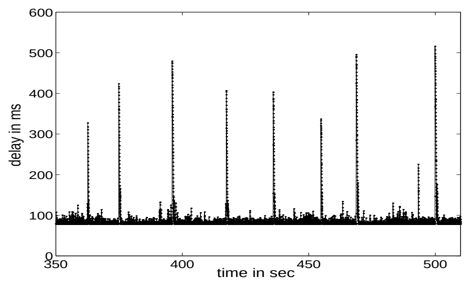


(b) Distance between consecutive spikes

Fig. 12. Statistics for the *random delay pattern*, THR-P1-ASH, Wed 2:00-3:00.



(a) Wed 0:00-1:00

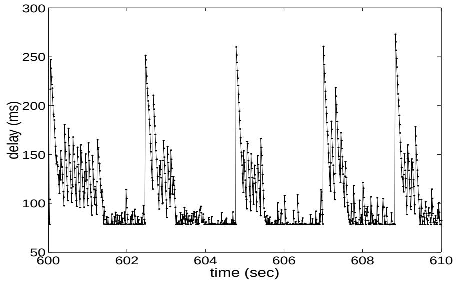


(b) Zooming in

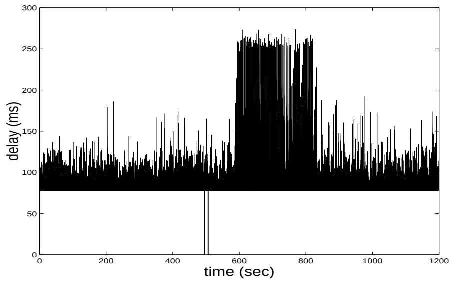
Fig. 13. Example of the *very high delay* pattern on path THU-P1-ASH.

spikes, we found that they all have the same structure.

Block Pattern. The second regular pattern consists of a cluster of spikes repeated periodically. The first spike is of a higher fixed height and is followed by many spikes half as high. Similar clusters of spikes are repeated periodically. Fig. 14 shows an example of this block pattern, which lasted for 5 minutes; the spikes were 250 ms high and the cluster was repeated every 2-3 sec. We call this pattern *the block pattern*, due to its box shape. It occurred 9 times in the entire measurement period

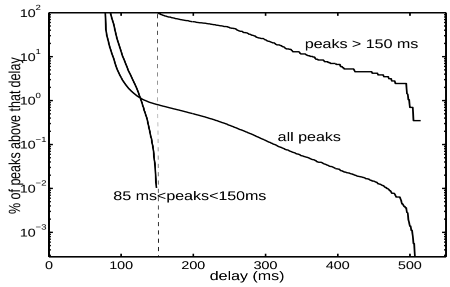


(a) 14:00-14:20 (UTC), Wed 06/27/01

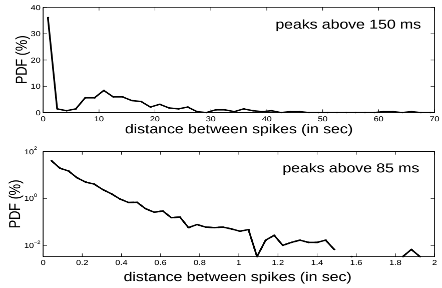


(b) Zooming in 10 seconds

Fig. 14. Example *block pattern* on THU-P1-ASH.



(a) CCDF of peaks of the spikes



(b) PDF of time between spikes

Fig. 15. Path THU-P_1-ASH. One hour (Wed 0:00-1:00) with random pattern and very high peaks.

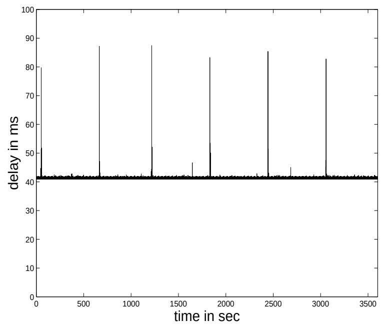


Fig. 16. Example of a path with very low delay variability (SJC-P7-ASH, Wed 4:00-5:00)

and it leads to the increase in the 50^{th} and 99^{th} delay percentiles in Fig. 9.

Delay Characterization in the Presence of Many Patterns. The normal delay pattern is the random one. In addition, when delays are high, one of the other patterns may also arise. In order to model such a trace we consider sets of peaks above a certain magnitude, starting with higher and proceeding with smaller magnitudes. We characterize every set of delays by describing the pattern and giving the statistics for how high the peaks are and how often they happen. One can then generate a set of peaks according to these distributions. The rest of probe delays can be generated from the peak delays, following the triangular spike shape. We find that lower delays usually follow the random pattern and have peaks roughly exponential (as in Fig. 12(a)) at roughly exponential distances (as in Fig. 12(b)); higher delays have more periodic structure.

As a concrete example, let us characterize the trace in Fig. 13(a) that has a random pattern and very high peaks. In Fig. 15, the CCDF of all peaks has a knee around 150 ms, and we choose to characterize separately the delays below and above 150 ms. In Fig. 15(a), we see that CCDF of peaks above 150 ms has a roughly exponential shape. In the top graph of Fig. 15(b), we see that the PDF of distances between these higher spikes has a maximum around 10 seconds, and can be as high as 70 sec. As for the lower delays, we see that the CCDF of their peaks is roughly a truncated exponential CCDF and the distance between them has a PDF with also an exponential shape (with a mean of 0.12 seconds or 12 packets). We model only spikes of significant size, i.e. above 85 ms, although 85% of all packet delays are in the [78 ms, 85 ms] range. If we considered all peaks, then the large majority of spikes would be small with distances of 1-2 packets from each other, thus hiding the higher spikes that are of interest to voice/video traffic.

Applying the same steps to every hour of the measurement period, we obtained similar distributions [7].

5.2.2 A Path with Very Low Delay Variability

Several of the measured paths had very low delay variability. For example, SJC-P_7-ASH is a path in a very well provisioned network that exhibits very low delay variations. Its delay percentiles were shown in Fig. 9(b), and an example of a perfectly triangular spike on this path was shown in Fig. 11(a). Fig. 16 shows an example hour on this path: delay is practically constant and 80ms high spikes happen once every 10 minutes.

We speculate that the 10 minutes periodicity is probably an artifact of the measurements collection itself: we collect and store probe measurements in a file every 10 minutes. Similar spikes in other traces, are probably hidden by additional spikes and larger delay variability.

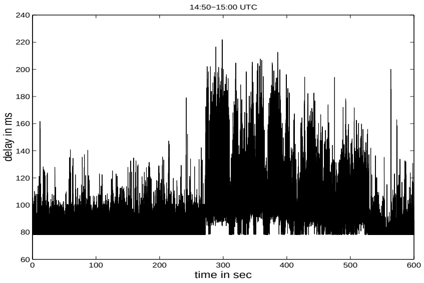


Fig. 17. Sustained increase in delay on THU-P1-ASH, Wed 06/27/01 14:50-15:00.

5.2.3 Low-Frequency Delay Components

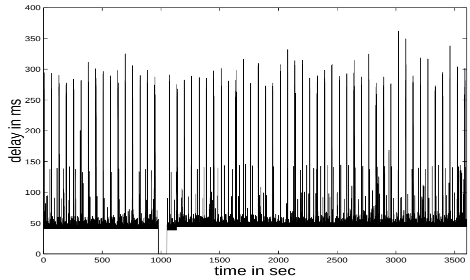
For most of the traces and for most of the time, delay jitter has the form of spikes that start from and return to the minimum delay. For these cases, the most natural characterization of delay is in terms of statistics for the height and the distance of spikes, as we did so far. In this section, we focus on the low-frequency delay components.

There are only a few cases where there is a sustained increase in delay. An example of a 10 minutes period with a sustained increase in delay, on a loaded path of provider P_1, is shown in Fig. 17(a). There is a sustained increase in delay lasting for hundreds of seconds. This example indicates that there is a low-frequency delay component, on top of which spikes are super-imposed. The same happened many times on this path (for example, this is the reason behind the increase in the median delay in Fig. 9(a)).

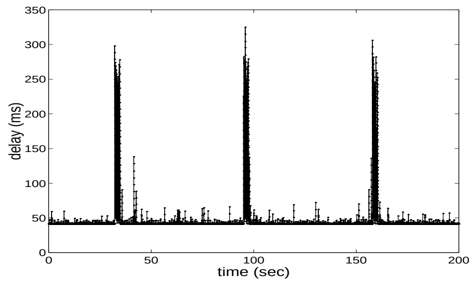
Mukherjee studied loss and delay on regional, backbone and cross-country paths, [12]. He sent probes infrequently (i.e. clusters of 20 probes, spaced 1 second apart, every 1 minute) to avoid increasing the load in the network. A spectral decomposition indicated the presence of dominant low-frequency delay components. He further smoothed out the average delay using a low-pass filter and found that the distribution of the smoothed delay was well approximated by a shifted gamma distribution.

We tested whether the delay in the example trace of Fig. 17, can be modeled using a gamma distribution. We found that Gamma distributions could fit well the smoothed delays, in several traces, as observed in their PDF, as well as in the accompanying quantile-quantile plot; we refer the interested user to [7] for details.

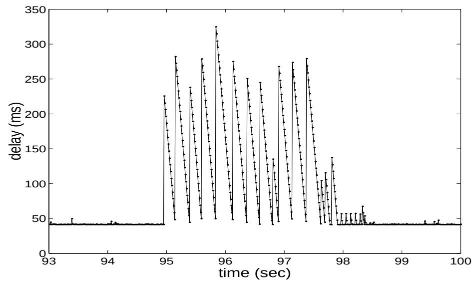
Unfortunately, the same was not true in several other cases. For example, the Gamma distribution was not a good fit for the smoothed delays in the right side of Fig. 17, where the slow-varying component is more pronounced. Furthermore, it was a bad fit when we considered the delays of all probes (instead of smoothed delays) or different averaging intervals and durations of the trace. Determining appropriate intervals for delay modeling, as well as the transitions between them, is a difficult problem, as it is also discussed in [19]. In conclusion, although the Gamma fit worked well for the low frequency delay component, this observation cannot be easily generalized.



(a) One hour: Wed, 21:00-22:00



(b) Zooming in 200 sec, starting at 21:10



(c) Zooming in 7 sec

Fig. 18. Periodic delay pattern on EWR-P_4-SJC.

5.2.4 Periodic Delay Patterns

We have already seen delay patterns that exhibit some structure. However, the most perfectly periodic pattern, was observed on all six paths of P_4, during the entire measurements period. Let us now consider the path $EWR - R4 - SJC$ and discuss it in detail. Fig. 9(d) showed that the maximum, 99.9^{th} and 99^{th} for this path have constantly high values for the entire day. A closer look reveals that these percentiles are due to the following periodic pattern.

Fig. 18(a) shows a typical hour on the path EWR-P4-SJC. Fig. 18(b) shows in detail 200 sec and Fig. 18(c) zoooms further in on 7 sec. We can see that there are clusters of spikes 250-300 ms high, lasting for 3 seconds and repeated every 60-70 sec. In addition, there are some smaller spikes (100-150 ms high). The periodicity in the high delay clusters is strikingly consistent.

To provide a characterization and allow the interested reader to reproduce the above delay pattern, we follow an approach similar to what we did for provider P_1. We first model higher delays (e.g. above 150 ms) and then we proceed with lower delays (e.g. above 100ms or above 50ms). The rationale is that special patterns happen for higher delays which are a small percentage of the entire data set and therefore they would get diluted if the entire data set were examined. Furthermore, the interested user can choose one of these sets, depending on the application (e.g. to simulate playout above a certain value). In [7], we considered we provided the distributions for the spike height and for the distance between spikes, for each one of the three sets (above 150ms, 100ms or 50ms). The statistics, omitted here for lack of space, capture the periodic clusters shown in Fig.19.

6 Discussion

6.1 *Effect on Applications*

Our study has been media-oriented in the collection (probes emulated VoIP packets) and as well as in the analysis phase (metrics of interest to stream traffic have been characterized). The interested reader is referred to [3,8] for a detailed discussion on how loss, delay and delay jitter affect the performance of multimedia traffic. In summary, most of the measured backbones exhibit good behavior for most of the time. However, there were paths consistently bad or periods of time (e.g. outages, complex loss events, or block and high delay patterns) when media traffic would perform poorly. During these periods, applications need support from the network and/or adaptive mechanisms at the end-systems (including playout scheduling, multi-path routing and rate-distortion optimized streaming [2]). Even TCP-based traffic could be affected by regular loss (section 4.4) and occasionally long round-trip times (section 5).

6.2 *Probable Causes*

Without privileged access to the network, we can only attempt to infer the causes behind our observations.

Single or a few packets lost are due to buffer drops. The regular drops on provider P_3 may be due to Random Early Drop (RED) turned on in the routers.

The longer *outage periods* accompanying changes in the minimum delay, can be attributed to routing changes and the time required by routing protocols to converge. For the reasons behind the rest of the outages, we speculate link failures or maintenance. Supporting evidence include the facts that (i) some outages happen at the same time of the day (that could be a maintenance process) and (ii) many outages affect more than one paths (implying failure of a shared link). Recent work in SprintLabs, [9], showed that the main problem in their backbone is link failures followed by periods of routing instability, during which packets are forwarded to invalid paths and eventually dropped.

We also observed *complex loss events*. Many of these clusters happened simultaneously on multiple paths and had the exact same loss pattern, hinting to a failure or congestion of a shared link.

In terms of *delay jitter*, we observed unusually high spikes, up to 500ms or 1sec. Spikes of smaller size can be due to multiplexing with cross traffic [1]. However, the perfect periodicity of the block patterns on provider P_1 (Section 5.2.1) and of the entire measurement period on 4 paths of the provider P_4 (see Section 5.2.4) cannot be explained by multiplexing with regular traffic. The size of these spikes and the lack of slow varying component in the delay traces, hints more toward "server vacations", i.e. periods during which routers do not serve packets to perform other internal tasks. This was recently observed in [13]; however the height of those spikes was much smaller than the ones we observed. Earlier experimental work [14] observed 600ms high spikes every 90 seconds, caused by a debugging option turned on in the gateways. They also identified other periodic patterns, [15], caused by synchronized routing updates due to faulty Ethernet interfaces. The periodicity and the height of the spikes we observed are more likely to be explained by network control traffic (such as exchanges of messages by routing protocols) or router specific operations (e.g. debugging options turned on, server vacations).

We observed *low correlation between increased loss and delay*, with a few exceptions that we highlighted when appropriate (e.g. outages correlated with changes in the fixed delay in sections 4.2, 5.1 and a few complex events correlated with packet drops in section 4.3).

Finally, we observed *consistent behavior of paths of the same provider*, i.e. only a limited combinations of loss and delay patterns happen on a particular path or group of paths, mainly depending on the provider. This is intuitively expected as these paths may share some network elements and operate under the provider's design and operation policies (e.g. regarding provisioning, network architecture, network protocols, maintenance activities etc).

6.3 *Using this study*

Our study can be used as input to various design problems concerned with adaptive mechanisms in the network and/or at the end-systems. Our characterization and modeling captures the behavior of backbones, which is an important part of the end-to-end path, especially for long-distance communication. Understanding this part of the end-to-end path is an important task in itself and/or in combination with other studies that characterize edge networks. Interestingly, and contrary to common belief, we identified several important problems introduced by backbone networks today and we provided insights into their causes. Paths at the edge introduce additional delay

Table 2
Summary of the events observed in the measurements, their probable causes, effect and possible remedies.

Impairment	Event observed	Possible Cause	Effect on VoIP	Possible Remedy
Loss	short duration	drop in the buffer	clipped speech	concealment
	loss clusters	reconfiguration,	loss of	improve network
	Outages	link failures	connectivity	reliability
Delay	high one-way	routing,	bad interactivity	live with it
	one-way delay	other components	amplified echo	cancel echo
Delay	high spikes	routers (debug, vacations)	clip, pitch change	fix the network,
Jitter	periodic spikes	control traffic	or extra delay	playout buffer

and loss, with their own characteristics. Applications eventually incur the superposition of all these effects.

Table 2 summarizes the observed impairments, their effect, probable causes, and remedies.

7 Conclusion

In this paper, we studied loss and delay measurements collected over backbone networks of major ISPs in the US. We used a rich data set which provides valuable insights into the behavior of Internet backbones. We described the loss and delay patterns observed in the measurements and we characterized their properties. We took a multimedia perspective, in that we sent probes that emulated stream traffic and we characterized properties that directly affect the performance of multimedia traffic. Most of the paths were found to exhibit fairly good characteristics while some others were found to introduce severe impairments to multimedia traffic. In the latter case, the causes of impairments seemed more related to the network operation (network protocols, failures and reconfiguration, router operation) rather than to congestion and traditional quality-of-service. We hope that our characterization of loss and delay will be used as input to the design and evaluation of various adaptive mechanisms.

Acknowledgment

We are grateful to RouteScience Technologies, now part of Avaya Inc., for providing the measurements.

References

[1] J.-C. Bolot, “Characterizing end-to-end packet Delay and Loss in the Internet”, in *Proc. ACM SIGCOMM 1993,* pp.289-298, San Francisco, USA, Sept. 1993.

[2] B. Girod, J. Chakareski, M. Kalman, Y. Liang, E. Setton, R. Zhang “Advances in Network-Adaptive Video Streaming”, *in Proc. IWDC 2002*, Capri, Italy, Sept. 2002.

[3] *ITU-T Rec. G.107*, “The Emodel, a computational model for use in transmission planning”, Dec. 1998.

[4] S.Jaiswal, G. Iannaccone, C. Diot, J. Kurose and D. Towsley (2004), “Inferring TCP Connection Characteristics Through Passive Measurements,” *in Proc. IEEE INFOCOM*, Hong Kong, March 2004.

[5] C. Labovitz, A. Ahuja, “Experimental Study of Internet Stability and Wide-Area Backbone Failures”, *in Proc. Fault-Tolerant Symp., FTCS99,* June 1999.

[6] D. Loguinov, H. Radha, “Performance of Low Bitrate Internet Video Streaming”, *in Proc. IEEE INFOCOM 2002,* New York, NY, USA, June 2002.

[7] A.Markopoulou,“Assessing the Quality of Multimedia Communications over Internet Backbone Networks”, *Ph.D. Thesis*, Stanford, Nov. 2002.

[8] A. Markopoulou, F. Tobagi, M. Karam, “Assessment of VoIP quality over Internet Backbones”, *in Proc. IEEE INFOCOM 2002,* New York, USA, June 2002.

[9] A.Markopoulou, G.Iannaccone, S.Bhattacharyya, C.N.Chuah, C.Diot, “Characterization of failures in an operational IP backbone”,*in Proc. IEEE INFOCOM 2004,* Hong Kong, March 2004.

[10] R. Mahajan, N. Spring, D. Wetherall, and T. Anderson, “User-level Internet Path Diagnosis”, *SOSP*, Oct. 2003.

[11] S. Moon, “Measurement and analysis of end-to-end delay and loss in the Internet”, *Ph.D. Thesis, UMASS,* Feb.’00.

[12] A. Mukherjee, “On the Dynamics and Significance of Low-Frequency Components in Internet Load”, *Internetworking: Research and Experience, Vol. 5, pp.163-205*, October 1994.

[13] K.Papagiannaki, S.Moon, C.Fraleigh, P.Thiran, F.Tobagi, C.Diot, “Analysis of Measured Single-Hop delay from Operational Backbone Network”, *in IEEE INFOCOM 2002,* New York, June 2002.

[14] D. Sanghi, A. Agrawala, B. Jain, “Experimental Assessment of end-to-end behavior on Internet”, *in IEEE INFOCOM 1993,* San Francisco, CA, USA, March 1993.

[15] D. Sanghi, O. Gudmundsson, A. Agrawala, “Study of network dynamics”, in *Proc. of 4th Joint European Network Conference 1993*, Norway, May 1993.

[16] N. Spring, R. Mahajan, D. Wetherall, “Measuring ISP Topologies with RocketFuel”, *in Proc. ACM SIGCOMM 2002*, Pittsburgh, PA, USA, August 2002

[17] F. Tobagi, A. Markopoulou, M.Karam, “Is the Internet ready for VoIP?”, *in Proc. IWDC’02*, Capri, Italy, Sept. 2002.

[18] M. Yajnik, S .Moon, J. Kurose, D. Towsley, “Packet Loss Correlation in the MBONE Multicast Network”, *UMASS Tech. Report, 1995.*

[19] Y. Zhang, N. Duffield, V. Paxson, S. Shenker, “On the Constancy of Internet Path Properties”, *Proc. of ACM SIGCOMM 2001,* San Diego, USA, Aug, 2001.