Impact of link failures on VoIP performance

Catherine Boutremans, Gianluca Iannaccone and Christophe Diot *

Abstract

We use active and passive traffic measurements to identify the issues involved in the deployment of a voice service over a tier-1 IP backbone network. Our findings indicate that no specific handling of voice packets (i.e. QoS differentiation) is needed in the current backbone but new protocols and mechanisms need to be introduced to provide a better protection against link failures. We discover that link failures may be followed by long periods of routing instability, during which packets can be dropped because forwarded along invalid paths. We also identify the need for a new family of quality of service mechanisms based on fast protection of traffic and high availability of the service rather than performance in terms of delay and loss.

1 Introduction

Recently, tier-1 Internet Service Providers (ISPs) have shown an ever increasing interest in providing voice and telephone services over their current Internet infrastructures. Voice-over-IP (VoIP) appears to be a very cost effective solution to provide alternative services to the traditional telephone networks.

However, ISPs need to provide a comparable quality both in terms of voice quality and availability of the service. We can identify three major causes of potential degradation of performance for telephone services over the Internet: network congestion, link failures and routing instabilities. Our goal is to study the frequency of these events and to assess their impact on VoIP performance.

We use passive monitoring of backbone links to evaluate the occurrence and impact of network congestion on data traffic. Passive measurements carried over different locations in the U.S. Sprint IP backbone allow us to study the transmission delay of voice packets and to evaluate the degree of congestion. However, this kind

*C. Boutremans (catherine.boutremans@epfl.ch) is with École Polytechnique Fédérale de Lausanne, Switzerland. G. Iannaccone (gianluca@sprintlabs.com) and C. Diot (cdiot@sprintlabs.com) are with Sprint ATL, Burlingame, CA.

of measurement cannot provide any information related to link failures or routing instabilities.

For this purpose, we have deployed an active measurement infrastructure in two locations well connected to the backbone. We capture and timestamp the probe packets at both ends to quantify losses and observe the impact of route changes on the voice traffic. We performed many week-long experiments in order to observe different link failure scenarios.

Given that all our measurements take place in the same Autonomous System (AS) we also complement our data with IS-IS routing information [7] collected in one of the backbone Points of Presence (POPs). This additional information give us a fairly complete view of the events that occur during our experiments. Indeed, active probes and routing information give us the capability of identifying precisely the links, the routers and even the interfaces that are responsible for failures or instabilities in the network.

Our findings indicate that the Sprint IP backbone network is ready to provide a toll-quality voice service. The level of congestion in the backbone is always negligible and has no impact on the voice quality.

On the other hand, link failures can impact the availability of VoIP services. We discovered that link failures may be followed by long periods of routing instability, during which packets can be dropped because forwarded along invalid paths. Such instabilities can last for tens of minutes resulting in the loss of reachability of a large set of end-hosts.

The paper is structured as follows. Section 2 briefly presents some related work, while Section 3 provides detailed information on the measurement approaches followed in this study. Section 4 describes the model used to assess the subjective quality of voice calls from transport level measurable quantities. In Section 5 we finally discuss our findings while Section 6 presents some concluding remarks.

2 Related work

Past literature on end-to-end Internet measurements has often focused on the study of network loss patterns and delay characteristics [5, 6, 14, 24, 22]. For example, Kostas [16] studied the feasibility of real-time voice over the Internet and discussed measured delay and loss characteristics. In order to evaluate the quality of Internet Telephony, [12] provided network performance data (in terms of delay and losses) collected from a wide range of geographically distributed sites. All these studies were based on round-trip delay measurements.

While information about delay and losses can give valuable insights about the quality of VoIP, they do not characterize the actual subjective quality experienced by VoIP users. In [9], Cole et al. propose a method for monitoring the quality of VoIP applications based upon a reduction of the E-model [2] to measurable transport level quantities (such as delay and losses).

Markopoulou et al. [17] use subjective quality measures (also based on the E-model) to assess the ability of Internet backbones to support voice communications. That work uses a collection of GPS synchronized packet traces. Their results indicate that some backbones are able to provide toll quality VoIP, today. In addition, they report that even good paths exhibit occasional long loss periods that could be attributed to routing changes. However, they do not investigate the causes of network failures neither the impact they have on the voice traffic.

3 Measurements

In this section we describe the two measurement approaches used in our study, i.e. the passive measurement system deployed in the Sprint IP backbone network and the active measurement system that uses probe packets to study routing protocols stability and link failures.

3.1 Passive measurements

The infrastructure developed to monitor the Sprint IP backbone consists of passive monitoring systems that collect packet traces on more than 30 links located in three POPs of the network. Details on the passive monitoring infrastructure can be found in [11].

In this study, we use traces collected from various OC-12 intra-POP links on July 24th, 2001, September 5th, 2001 and November 8th, 2001. A packet trace contains the first 44 bytes of every IP packet that traverses the monitored link. Every packet record is also timestamped using a GPS reference signal to synchronize timing information on different systems [18].

We use the technique described in [21] to compute one-way delays across the Sprint backbone. The basic idea behind that technique is to identify those packets that enter the Sprint backbone in one of the monitored POPs and leave the network in another one. Once such packets are identified computing the delays simply requires to compute the difference between the recorded timestamps.

3.2 Active measurements

Passive measurements provide valuable information about network characteristics, but the data collected depend on the traffic generated by other parties, which is completely out of our control. Moreover, given that we do not monitor all the links

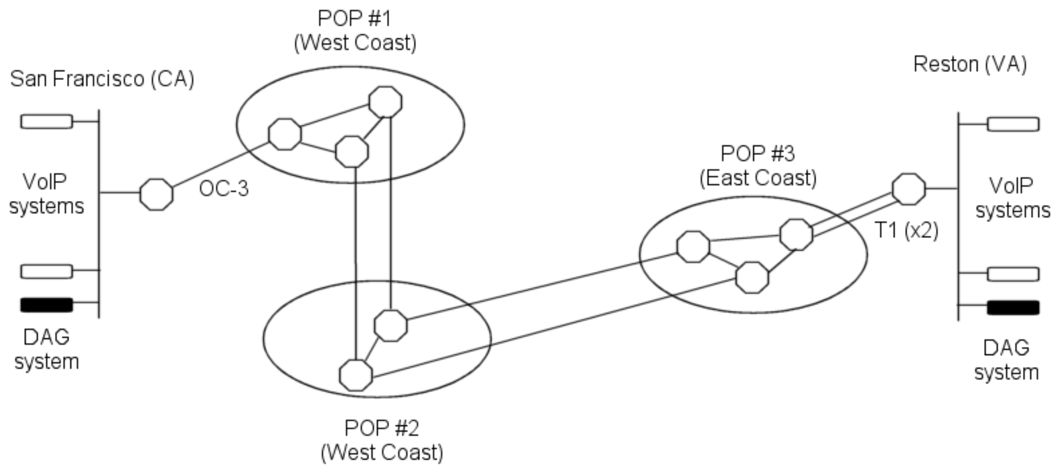

Figure 1: Architecture and connection of the active measurement systems.

of the backbone network, we are not able to measure jitter or loss rates through simple passive monitoring (packets may leave the network through not monitored links) [21]. Therefore, our passive measurements alone cannot provide results on the quality of the voice calls. These are the motivations behind the use of active measurements to complement the passive ones. In an active measurement environment we can perfectly control the amount and the characteristics of the traffic that we inject in the network and thus draw precise conclusions about the impact of the network on the monitored traffic.

3.2.1 Measurement infrastructure

We deployed active measurement systems in two locations of the U.S. (Reston, VA and San Francisco, CA) well connected to the Sprint backbone, i.e. just one router away from the backbone network. Figure 1 shows the architecture of the testbed and the way the sites are connected through the Sprint network (the thick lines indicate the path followed by our traffic). Note that each access router in a POP is connected to two backbone routers for reliability and, usually, per-destination prefix load balancing is implemented.

The access links to the backbone were chosen to be unloaded in order not to introduce additional delay. At the end of each experiment we verified that no packet losses were induced on the last hops of the paths.

In each site, four systems running FreeBSD generate a traffic made of 200 byte UDP packets at a constant rate of 50 packets per second. We choose this rate so that the probes could be easily used to emulate a voice call compliant to the G.711 standard [1].

An additional system captures and timestamps the probe packets using a DAG3.2e card [8]. The DAG cards provide very accurate timestamping of packets synchronized using a GPS (or CDMA) receiver [18]. The probe packets are recorded and timestamped right before the access links of the two locations in both directions.

In the experiment we discuss here, probes are sent from Reston (VA) to San Francisco (CA) for a duration of 2.5 days starting at 04.00 UTC on November 27th, 2001. We have run active measurements for several weeks but we have chosen that specific trace because it exhibits an interesting network failure event. In terms of delay, loss and voice call quality we have not measured, instead, any significant difference among the many different experiments.

3.2.2 Routing data

We integrate our measurement data with IS-IS routing information collected in POP#2 (see Figure 1). We use an IS-IS listener [19] to record all routing messages exchanged during the experiment. IS-IS messages permit to correlate loss and delay events to changes in the routing information. In order to illustrate the kind of data that are collected by the listener, we give a brief description of the IS-IS protocol.

IS-IS [20] is a link state routing protocol used for intra-domain routing. With IS-IS, each link in the network is assigned a metric value (weight). Every router[1] broadcasts information about its direct connectivity to other routers. This information is conveyed in messages called Link State PDUs (LSP). Each LSP contains information about the identity and the metric value of the adjacencies of the router that originated the LSP. In general, a router generates and transmits its LSPs periodically, but LSPs are also generated whenever the network topology changes (e.g. when a link or a router goes up or down). Thus, LSPs provide valuable information about the occurrence of events such as loss of connectivity, route changes, etc.

Once a router has received path information from all other routers, it constructs its forwarding database using Dijkstra's Shortest Path First (SPF) algorithm to determine the best route to each destination. This operation is called the decision process. In some transitory conditions (e.g. after rebooting), the decision process can take a considerable amount of time (several minutes) since it requires all the LSPs to be received in order to complete. During that transitory period, a router is responsible to make sure that other routers in the network do not forward packets towards itself. In order to do so, a router will generate and flood its own LSPs with

[1] IS-IS has been designed within the ISO-OSI standardization effort using the OSI terminology. In this paper, we have instead decided to avoid the use of OSI terms.

the "Infinite Hippity Cost" bit set[2]. This way, other routers will not consider it as a valid node in the forwarding paths.

4 Voice call rating

Even though active measurements may provide accurate information on network delay and losses, such statistics are not always appropriate to infer the quality of voice calls. In addition to measurements, we use a methodology to emulate voice calls from our packet traces and assess their quality using the E-model standard [2, 3, 4].

4.1 A voice quality measure: the E-model

The E-model predicts the subjective quality that will be experienced by an average listener combining the impairment caused by transmission parameters (such as loss and delay) into a single rating. The rating can then be used to predict subjective user reactions, such as the Mean Opinion Score (MOS). According to ITU-T Recommendation G.107, every rating value corresponds to a speech transmission category, as shown in Table 1. A rating below 60 indicates unacceptable quality, while values above 70 correspond to PSTN quality (values above 90 corresponding to very good quality).

R-value range	MOS	Speech transmission quality
$100 - 90$	4.50-4.34	best
$90 - 80$	4.34-4.03	high
$80 - 70$	4.03-3.60	medium
$70 - 60$	3.60-3.10	low
$60 - 0$	3.10-1.00	very poor

Table 1: Speech transmission quality classes and corresponding rating value ranges.

The E-model rating R is given by:

$$R = R_0 - I_s - I_d - I_e + A \tag{1}$$

where R_0 groups the effects of noise, I_s represents impairment that occur simultaneously with the voice signal (quantization), I_d is the impairment associated

[2] This bit is also referred to as the OverLoad (OL) bit.

with the mouth-to-ear delay, and I_e is the impairment associated with signal distortion (caused by low bit rate codecs and packet losses). The advantage factor A is the deterioration that callers are willing to tolerate because of the 'access advantage' that certain systems have over traditional wire-bound telephony, e.g. the advantage factor for mobile telephony is assumed to be 10. Since no agreement has been reached for the case of VoIP services, we will drop the advantage factor in this study.

4.2 Reduction of the E-model to transport level quantities

Although an analytical expression for I_d is given in [3] and values for I_e are provided in Appendix A of [4] for different loss conditions, those standards do not give a fully analytical expression for the R-factor. In this work, we use a simplified analytic expression for the R-factor that was proposed in [9] and that describes the R-factor as a function of observable transport level quantities.

In this section, we briefly describe the reduction of equation (1) to transport level quantities as proposed in [9] and we introduce the assumptions made about the VoIP connections under study.

4.2.1 Signal-to-noise impairment factors R_0 and I_s

Both R_0 (effect of background and circuit noise) and I_s (effect of quantization) describe impairment that have to do with the signal itself. Since none of them depend on the underlying transport network, we rely upon the set of default values that are recommended in [3] for these parameters. Choosing these defaults values, the rating R can be reformulated as:

$$R = 94.2 - I_d - I_e \tag{2}$$

4.2.2 Delay impairment I_d

ITU-T Recommendation G.107 [3] gives a fully analytical expression for I_d in terms of various delay measures (such as mouth-to-ear delay, delay from the receive side to the point where signal coupling occurs and delay in the four wire loop) and other parameters describing various circuit switched and packet switch inter-working scenarios.

Since we focus, in this work, on pure VoIP scenarios, we make the following simplifications: i) the various delay measures collapse into a single one, the mouth-to-ear delay, and, ii) the default values proposed in [3] are used for all parameters in the expression of I_d other than the delay itself. In particular, the influence of echo is supposed negligible. The curve obtained describing I_d as a function of the mouth-to-ear delay can then be approximated by a piece-wise linear function [9]:

$$I_d = 0.024d + 0.11(d - 177.3)H(d - 177.3) \tag{3}$$

where d is the mouth-to-ear delay and H is the Heavyside function. d is composed of the encoding delay (algorithmic and packetization delay), the network delay (transmission, propagation and queuing delay) and the playout delay introduced by the playout buffer in order to cope with delay variations. The Heavyside function is defined as follows:

$$\begin{aligned} H(x) &= 0 \quad \text{if } x < 0 \\ H(x) &= 1 \quad \text{if } x \geq 0 \end{aligned} \tag{4}$$

4.2.3 Equipment impairment I_e

The impairment introduced by distortion are brought together in I_e. Currently, no analytical expression allows to compute I_e as a function of parameters such as the encoding rate or the packet loss rate. Estimates for I_e must be obtained through subjective measurements. A few values for I_e are given in Appendix A of [4] for several codecs (i.e. G.711, G.729,...) and several packet loss conditions.

In this work, we focus on the G.711 coder which does not introduce any distortion due to compression (and hence leads to the smallest equipment impairment value in absence of losses). In addition, we assume that the G.711 coder in use implements a packet loss concealment algorithm. In these conditions, the evolution of the equipment impairment factor I_e as a function of the average packet loss rate can be well approximated by a logarithmic function. In particular, if we assume that we are in presence of random losses, the equipment impairment can be expressed as follows [9]:

$$I_e = 30 \ln(1 + 15e) \tag{5}$$

where e is the total loss probability (i.e., it encompasses the losses in the network and the losses due to the arrival of a packet after its playout time).

In summary, the following expression will be used to compute the R-factor as a function of observable transport quantities:

$$\begin{aligned} R \;=\;& 94.2 - 0.11(d - 177.3)H(d - 177.3) - \\ & - 0.024d - 30 \ln(1 + 15e) \end{aligned} \tag{6}$$

where d is the mouth-to-ear delay, e is the total loss probability and H is the Heavyside function defined in equation (4).

4.3 Call generation and rating

In order to assess the quality of voice calls placed at random times during the measurement period, we emulate the arrival of short business calls. We pick call arrival times according to a Poisson process with a mean inter-arrival time of 60 seconds. We draw the call durations according to an exponential distribution with a mean of 3.5 minutes [15]. The randomly generated calls are then applied to the packet traces for quality assessment.

Since IP telephony applications often use silence suppression to reduce their sending rate, we simulate talkspurt and silence periods within each voice call using for both periods an exponential distribution with an average of 1.5s [13]. Packets belonging to a silence period are simply ignored.

At the receiver end, we assume that a playout buffer is used to absorb the delay variations in the network. The playout delay is defined as the difference between the arrival and the playout time of the first packet of a talkspurt. Within a talkspurt, the playout times of the subsequent packets are scheduled at regular intervals following the playout time of the first one. Packets arriving after their playout time are considered lost. A playout buffer can operate in a fixed or an adaptive mode. In a fixed mode, the playout delay is always constant while in an adaptive mode, it can be adjusted between talkspurts.

In this work, we opt for the fixed playout strategy because the measured delays and jitters are very small and a fixed playout strategy would represent a worst case scenario. Thus, we implement a fixed playout delay of 75ms (which is quite high, but still leads to excellent results, as described in Section 5).

The quality of the calls described above is then computed as follows. For each talkspurt within a call, we compute the number of packet losses in the network and in the playback buffer. From these statistics, we deduce the total packet loss rate e for each talkspurt. In addition, we measure the mouth-to-ear delay d, which is the sum of the packetization delay (20ms, in our case), the network delay of the first packet of the talkspurt and the playout delay.

In order to assess the quality of a call we apply equation (6) to each talkspurt and then we define the rating of a call as the average of the ratings of all its talkspurts.

5 Results

In this section we discuss our findings derived from the experiments and measurements. We first compare the results obtained via the passive and active measurements and then focus on the impact of link failures on traffic. We conclude with a discussion of the call rating using the methodology proposed in Section 4.

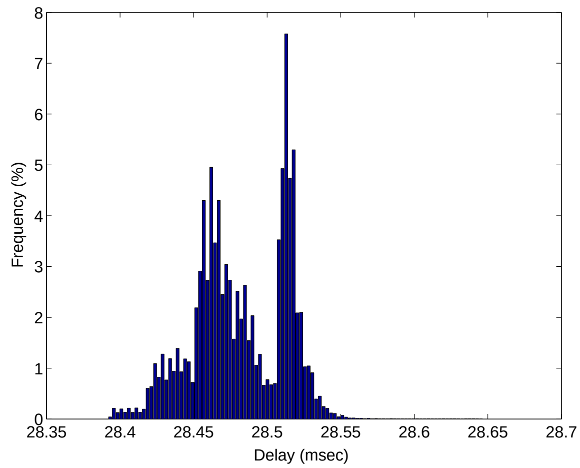

Figure 2: Passive measurements: distribution of the one-way transmission delay between East and West Coast of the U.S.

5.1 Delay measurements

In Figure 2 we show the one-way delay between two Sprint POPs on the East and West Coast of the United States. The data shown refers to a trace collected from the passive measurement system on July 24th 2001, however we have systematically observed similar delay distributions on all the traces collected in the Sprint IP monitoring project [11]. The delay between the two POPs is around 28.50ms with a maximum delay variation of less than 200μs. Such delay figures show that packets experience almost no queueing delay and that the element that dominates the transmission delay is the propagation over the optical fiber [21].

We performed the same delay measurements on the UDP packets sent every 20ms from Reston (VA) to San Francisco (CA) for a period of 2.5 days. Figure 3 shows the distribution of the one-way transmission delay. The minimum delay is 30.95ms, the average delay is 31.38ms while the 99.9% of the probes experience a delay below 32.85ms.

As we can see from the figures, the results obtained by the active measurements are consistent with the ones derived from passive measurements. Low delays are a direct result of the over-provisioning design strategies followed by most tier-1 ISPs in the attempt to keep the maximum link utilization for backbone links below the threshold of 50%. Such strategy is dictated by the need for commercial ISPs to be highly resilient to network failures and always capable of handling short-term variations in the traffic demands.

The delay distribution in Figure 3 shows also another interesting feature: an

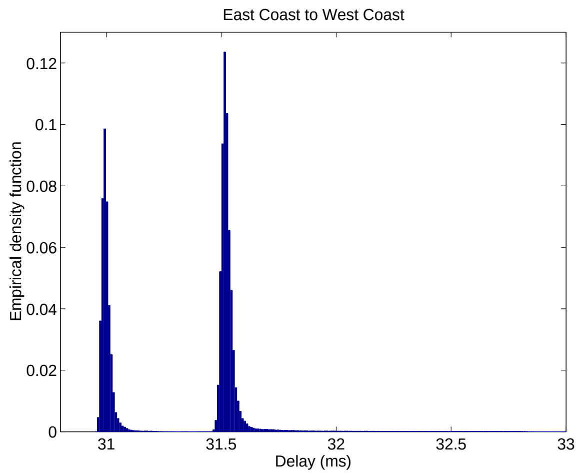

Figure 3: Active measurements: distribution of the one-way transmission delay from Reston (VA) to San Francisco (CA).

event of re-routing has occurred during the experiment. The distribution shows two spikes that do not overlap and for which we can thus identify two minima (30.96ms and 31.46ms), that represent the propagation delays of the two routes[3].

While the difference between the two minima is relatively high (around 500μs), the difference in router hops is just one (derived from the TTL values found in the IP packets). One additional router along the path does not justify the 500μs difference [21], thus the major component of the delay difference is the length of the two fiber paths (500μs correspond roughly to 100km). This is another typical characteristic of the design of backbone networks: between each pair of POPs there are at least two IP routes that use two physically disjoint fiber paths.

5.2 Impact of failures on data traffic

In this section we investigate further the re-routing event. To the best of our knowledge there is no experimental study on failures and their impact on traffic on an operational IP backbone network. The difficulties involved in collecting data on the traffic at the time of a failure is probably the main reason for that. Within several weeks of experiments, our VoIP traffic has suffered a single failure. Nevertheless, we believe it is fundamental for researchers and practitioners to study such

[3]The delay distribution derived from passive measurements shows also some spikes. In that case, however, we cannot distinguish between different delays due to packet sizes [21] or due to routing, given that we do not dispose of the routing information.

failure events in order to validate the behaviors and performance of routing protocols, routing equipment and to identify appropriate traffic engineering practices to deal with failures.

The failure perturbed the traffic during a 50 minutes period between 06:30 and 07:20 UTC on November 28th, 2002. During that failure event, the traffic experienced various periods of 100% losses before being re-routed for the rest of the experiment (33 hours).

We now provide an in-depth analysis of the series of events related to the failure and of the causes of loss periods. In addition to active measurements, we use the routing data collected by our IS-IS listener.

Figure 4 shows the delay that our voice packets experience at the time of the failure. Each dot in the plot represents the average delay over a five-second interval. Figure 5 provides the average packet loss rate over the same five-second intervals.

At time 06:34, a link failure is detected and packets are re-routed along an alternative path that results in a longer delay. It takes about 100ms to complete the re-routing during which all the packets sent are lost. Although the quality of a voice call would certainly be affected by the loss of 100ms worth of traffic, the total impact on the voice traffic is minimal given the short time needed for the re-routing and the small jitter induced (about 500μs).

After about a minute, the original route is restored. A series of 100% loss periods follows, each of which lasting several seconds. Figure 6 shows the one-way delay experienced by all the packets during one of such 100% loss periods (the same behavior can be observed in all the other periods). As we can see from the figure, no packets are buffered by the routers during the short outages (no packets experience long delays) but they are just dropped because forwarded along an invalid path. Figure 7 shows the sequence numbers of the packets as received by the end host on the West Coast. Again, no losses nor re-orderings occur during those periods. This is a clear indication that packet drops are not due to congestion events but due to some kind of interface or router failure.

At time 06:48, the traffic experiences 100% losses for a period of about 12 minutes. Surprisingly, during that period no alternative path is identified for the voice traffic. At time 07:02 a secondary path is found but there are still some 100% loss periods. Finally, at 07:19, the original path is operational again and at time 07:36 an alternative path is chosen and used for the remaining part of the experiment.

Until now we have limited the description of the failure to what can be observed via end-to-end measurements. In the following we discuss our findings derived from the analysis of the routing data. Figure 8 illustrates the portion of the network topology with the routers involved by the failure. The routers (R_1 to R_5) are located in 2 different POPs. The solid arrows show the primary path followed by the traffic.

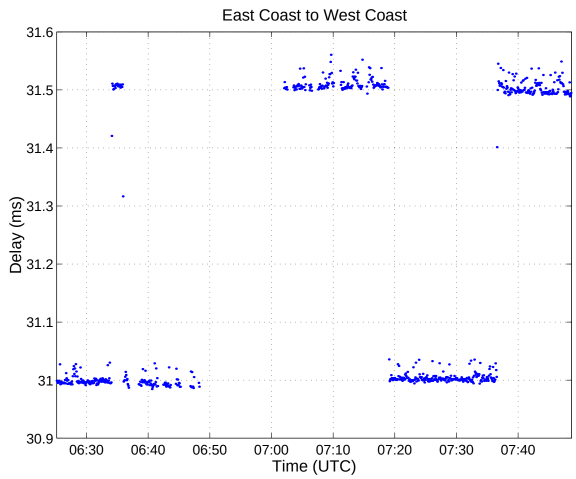

Figure 4: Average delay during the failure. Each dot corresponds to a five-second interval

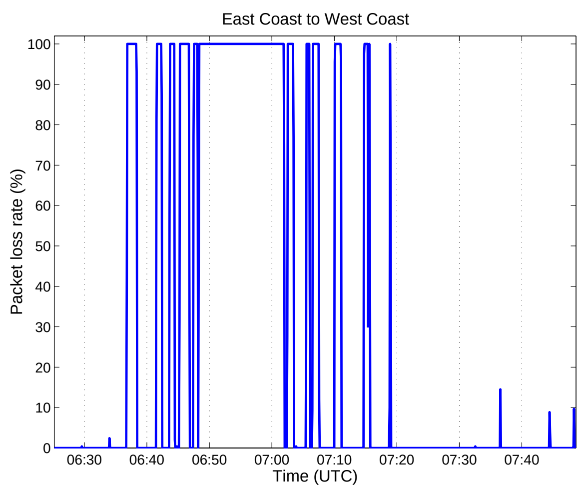

Figure 5: Average packet loss rate during the failure computed over five-second intervals

The dashed arrows show the alternative path used after the failure.

Table 2 summarizes all the messages that we have collected from the IS-IS listener during the experiment and their impact on traffic. The "Time" column indicates the time at which the LSPs are received by our listener, the central column ("IS-IS LSPs") describes the LSPs in the format <senders>:<content>, while the

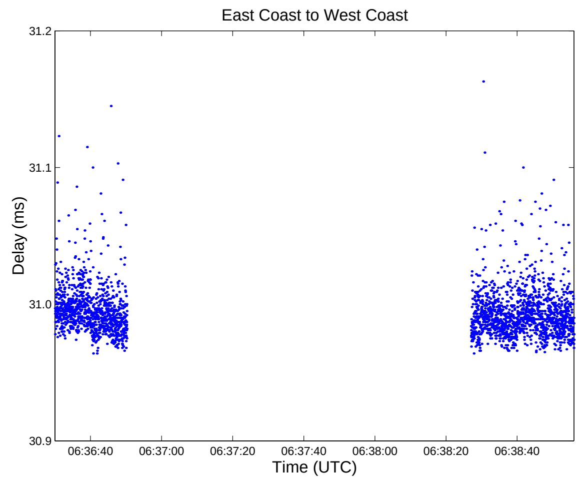

Figure 6: One-way delay of voice packets during the first 100% loss period

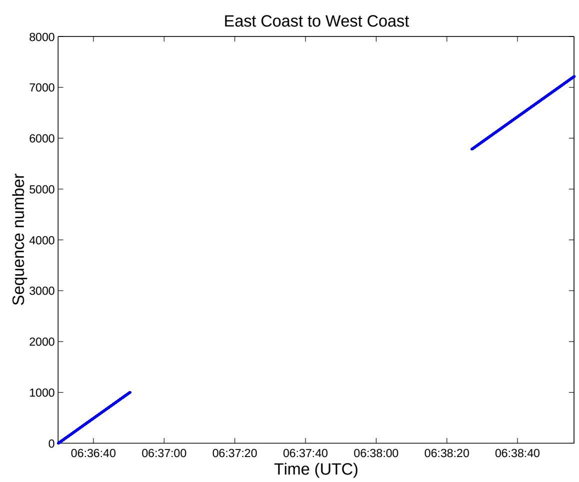

Figure 7: Sequence numbers of received voice packets during the first 100% loss period

third column describes the impact on the traffic of the event reported by IS-IS.

At the time of the first re-routing, routers R_1, R_2 and R_5 report via IS-IS the *loss of adjacency* with R_4. The fact that all the links from R_4 are signaled down represents a strong indication that this failure is a router failure as opposed to link failure. As we said earlier, the network reacts to this first event as expected. In about 100ms, R_5 routes the traffic along the alternative path through R_2 (i.e. R_5-

R_3-R_2-R_1).

Time	IS-IS LSPs	Impact on traffic
06:34	R_1, R_2, R_5: link to R_4 is down	Re-routed through R_3 in 100ms
06:35	R_1, R_2, R_5: adjacency with R_4 recovered	Re-routed through R_4
from 06:59 to 07:06	R_1: link to R_4 "flaps" 7 times	100% loss periods. Re-routed through R_3
from 07:00 to 07:17	R_2: link to R_4 "flaps" 5 times	100% loss periods. Re-routed through R_3
from 07:04 to 07:17	R_5: link to R_4 "flaps" 4 times	100% loss periods. Re-routed through R_3
07:07	R_1: link to R_4 is down	Re-routed through R_2
07:17	R_1, R_2, R_5: link to R_4 is definitely up	Traffic restored on the original path

Table 2: Summary of the events occurred during the network outage.

In the period from 06:35 to 06:59, the IS-IS listener receives several (periodic) LSPs from all the five routers reporting that all the links are fully operational. During that time, though, the traffic suffers from 100% loss periods; a router failure results, indeed, in unstable (and unpredictable) behaviors. For about 13 minutes, R_4 oscillates between a normal operational state (i.e. it forwards the packets without loss or additional delay) and a "faulty" state during which all the traffic is dropped. However, such "faulty" states never last long enough to give a chance to

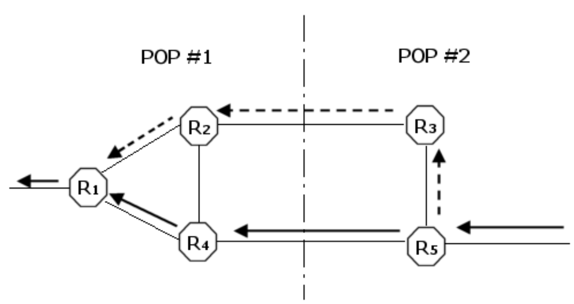

Figure 8: Routers involved by the failure. The solid arrows indicate the primary path for our traffic. The dashed arrows indicate the alternative path through R_3.

the other routers to detect the failure.

At time 06:48, R_4 finally reboots. It then starts collecting LSP messages from all the routers in the network in order to build its own routing table. This operation is usually very expensive for a network of the size of the Sprint IP backbone. It may require minutes to complete: the router has to collect the LSP messages that all the other routers periodically send.

While collecting the routing information, R_4 does not have a routing table and is therefore not capable of handling any packet. As we described in Section 3, a router is expected to send LSP messages with the “Infinity Hippity Cost” bit set. In our case, instead, R_4 does not set that bit and, thus, R_5 forwards the voice traffic to R_4 where it is dropped for 12 minutes.

At time 7:02, R_4 builds its first routing table and the traffic is partially restored but the links R_2-R_4 and R_5-R_4 start flapping resulting again in 100% loss periods. Note that the traffic is only restored along the alternative path (hence, the longer delays) because the link between R_1 and R_4 is reported to be down. We conjecture that the 100% loss periods are due to R_5 forwarding traffic to R_4 every time the link $R_4 - R_5$ is up, although R_4 does not have a route to R_1.

Most likely the links are not flapping because of an hardware problem but because R_4 is starting receiving the first BGP updates[4] force frequent re-computations of the routing table to add new destination prefixes.

Finally, at time 07:19 all routers report that the links with R_4 are up and the routing remains stable for the rest of the experiment. Traffic is however re-routed again along the alternative path after about 18 minutes even if the original path is operational. This is simply due to the fact that R_5 modifies its load balancing policy over the two equal cost paths (solid and dashed arrows in Figure 8). Routers that perform per-destination prefix load balancing (as R_5 in our case) can periodically modify their criteria (i.e., which flow follow which path) in order to avoid that specific traffic patterns defeat the load balancing (e.g., if most of the packets belong to few destination prefixes, one path may result more utilized than the other).

In order to summarize our findings, we can divide the failure we observed in two phases:

- The first phase from time 06:34 to 06:59 is characterized by instabilities in the packet forwarding on router R_4: only few LSPs are generated but the traffic experience periods of 100% packet loss. Such “unstable” phase is due to the particular type of failure that involved an entire router and most likely the operating system of the router. The effect on packet forwarding and routing is thus unpredictable and difficult to control protocol-wise.

[4] R_4 can setup the I-BGP sessions, that run over TCP, with its peers only once it has a valid routing table, i.e. it has received all LSP updates.

- The second phase goes from time 06:48 to 07:19 and is instead characterized by a very long outage followed by some routing instabilities and periods of 100% loss. We have not been able to identify the reason why R_4 did not set the "Infinity Hippity Cost" bit, however there are two possibilities: i) a bug in the IS-IS implementation or ii) a misconfiguration in router R_4.

It is important to observe that both the first and the second phase of the failure event are not due to the IS-IS routing protocol. Therefore, we do not expect that the use of a different routing protocol (e.g. "circuit-based" routing mechanisms such as MPLS [23]) would mitigate the impact of failures on traffic.

Instead, it is our opinion that router vendors and ISPs should focus their efforts on the improvement of the reliability of routing equipment, intended both in terms of better hardware architectures and more stable software implementations. Another important direction of improvement is certainly the introduction of automatic validation tools for routers' configurations. However, such tools would require first to simplify the procedures to configure the routing protocols in order to be able to completely understand the interactions among the configuration parameters. We do not believe that introducing circuits or label-switched paths will help in such simplification effort.

5.3 Voice quality

In this paper, we investigate two different quality metrics characterizing a VoIP service: i) the availability of the service, and ii) the quality of voice calls when the service is available. This section is devoted to the study of the latter, namely the quality experienced by a VoIP user who was able to place a call.

In order to evaluate the quality of the voice calls we emulated a set of voice calls using the delay and loss statistics of the probes as described in Section 4. Figure 9 shows the rating of the voice calls during the 2.5 days of the experiment. We did not place any call during the failure event (50 minutes out of the 2.5 days) because we are interested in the quality of completed calls.

Figure 10 shows the distribution of call quality for the 2.5 days of experiment. All these results were derived assuming a fixed playout buffer. One can notice that the quality of calls does not deviate much from its mean value which is fairly high: 90.27. Among the 3,364 calls that were placed, only one experiences a quality rating below 70, the lower threshold for toll-quality. We are currently in the process of investigating what caused the low quality of some calls. Moreover, with 99% of calls experiencing a quality above 84.68, our results confirm that the Sprint IP backbone can support a voice service with toll quality standards.

The very good quality of voice traffic is a direct consequence of the low delays, jitter and loss rates that probes experience. Without taking into account the 50

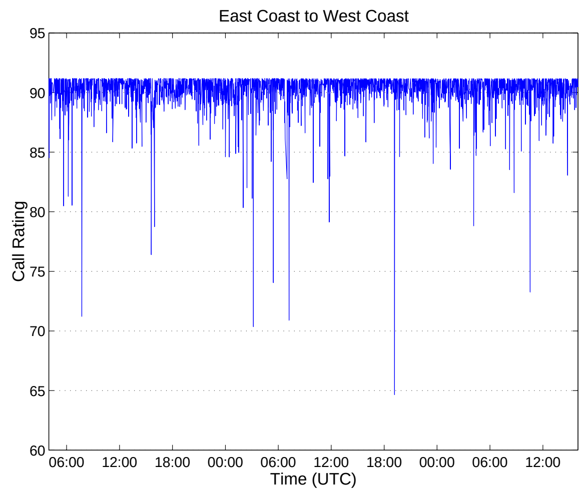

Figure 9: Voice call ratings (excluding the failure event)

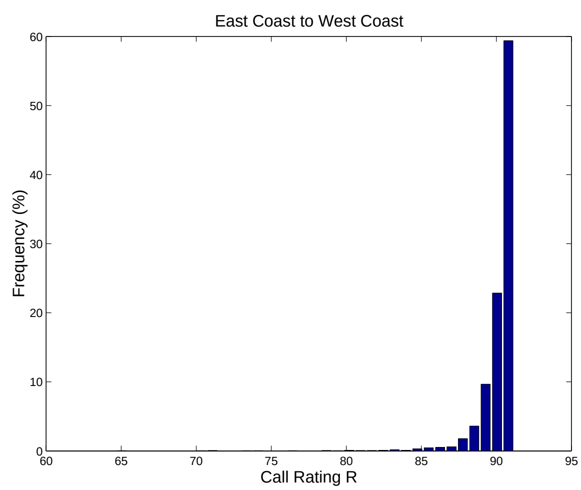

Figure 10: Distribution of voice call ratings (excluding the failure event)

Burst length	Frequency of occurence
1	90.42%
2	8.71%
3	0.71%
4 and above	0.16%

Table 3: Repartition of burst lengths (excluding the failure event)

minutes of failure, the average loss rate is 0.19%. Even if we count in the failure, the loss rate over the 2.5 days of the experiment is equal to 1.15%.

We also studied the probability of having long bursts of losses. Our goal is to verify that the assumptions on the distribution of packet losses (we assumed that the losses were not bursty) and on the performance of packet loss concealment techniques are well suited to our experiment.

For this purpose, we define the *loss burst length* as the number of packets dropped between two packets correctly received by our end hosts. Table 3 shows the repartition of burst length among the losses observed during the period of experiment. The vast majority (90.42%) of loss events have a burst length 1, while 99.84% of the events have a burst length less than 4. This tends to indicate that the packet loss process is not bursty and hence validates the hypothesis of "random losses" that was put forward in Section 4. Moreover, with a large majority of isolated losses, we can conjecture that packet loss concealment techniques would be efficient in attenuating the impact of packet losses.

6 Conclusion

We have presented the results of a set of active and passive measurements carried over the Sprint IP backbone network. We have run experiments for several weeks and we can derive the following conclusions:

- A voice service based on VoIP is certainly feasible over the Sprint IP backbone network. Delay and loss figures indicate that the quality of the voice calls would be comparable with that of traditional telephone networks. We have pointed out that voice quality is not the only metric of interest for evaluating the feasibility of a VoIP service. The availability of the service also covers a fundamental role and it will represent a major issue in the deployment of VoIP over commercial ISPs.

- The major causes of quality degradation are currently link and router failures. We have observed that despite careful IP route protection, link failures can significantly impact an IP voice service. Moreover, given the high operational costs involved in maintaining alternative backup paths for all the traffic, we foresee the need for a new family of quality of service mechanisms based on protection. In this case, customers could decide to protect against routing failures only portions of their traffic.

- The reliability of routing equipment represents a major obstacle to the introduction of VoIP services. In the past years, the exponential increase in traffic demands and in the number of hosts connected to the Internet has made extremely difficult to address such reliability issues. Nevertheless, ISPs willing to provide a voice service in addition to traditional best-effort data services need to identify valid metrics for the reliability of the routing hardware equipment and, in particular, of software implementations.

- We foresee the need for automatic verification tools of routers' configurations. The Sprint IP backbone network consists today of more than a thousand routers and it is still growing at a very fast pace. Clearly, it is very difficult to manually validate that the configurations are consistent throughout the entire network. Further investigation is needed to identify all the interactions between the various protocols (e.g. IS-IS, I-BGP and E-BGP) in order to implement such tools.

- The introduction of circuit- or label-switched networks will not help in mitigating the impact of failures. The failure event we have described in Section 5 is a clear example of this. As long as the failure is reported in a consistent way by the routers in the network, the IS-IS protocol can efficiently identify alternative routes (the first re-routing required only 100ms to complete). Such recovery times are the same ones provided, for example, by the MPLS Fast-ReRoute (FRR) mechanism introduced by Cisco [10]. However, a failing and unstable router that sends invalid messages would cause also MPLS (or any other routing protocol) to fail.

Future work will involve more experiments. Through long-term measurements, we aim to evaluate the likelihood of link and node failures in a tier-1 IP backbone. We also intend to address the problem of VoIP traffic traversing multiple autonomous systems.

Another important area will be the study of metrics that permit to compare the telephone network reliability with the Internet reliability. On telephone networks, the most common metrics are based on the downtime of individual switches or access lines. The objective of such metrics is to measure the impact on customers of network outages. The Federal Communications Commission requires telephone operators to report any outage that affects more that 90,000 lines for at least 30 minutes. Such metrics are however difficult to apply to the Internet for a few reasons: i) there is no definition of "line" that can be applied to the Internet; ii) it is very difficult to count how many customers have been affected by a failure; iii) from a customer standpoint there is no difference between outages due to the network or due to the servers (e.g. DNS servers, web servers, etc.).

References

[1] Pulse code modulation (PCM) of voice frequencies. ITU-T Recommendation G.711, November 1988.

[2] Speech processing, transmission and quality aspects (STQ); overall transmission plan aspects for telephony in private networks. ETSI Guide 201 050, February 1999.

[3] The e-model, a computational model for use in transmission plannning. ITU-T Recommendation G.107, May 2000.

[4] Provisional planning values for equipment impairment factor Ie. Appendix to ITU-T Recommendation G.113, February 2001.

[5] J.-C. Bolot. End-to-end packet delay and loss behavior in the internet. In *Proc. ACM Sigcomm'93*, pages 189–199, August 1993. San Francisco, CA.

[6] M. Borella. Measurement and interpretation of internet packet loss. *Journal of Communication and Networks*, 2(2):93–102, June 2000.

[7] R. Callon. Use of OSI IS-IS for routing in TCP/IP and dual environments. RFC 1195, Dec. 1990.

[8] J. Cleary, S. Donnelly, I. Graham, A. McGregor, and M. Pearson. Design principles for accurate passive measurements. In *Proceedings of Passive and Active Measurement Workshop*, Apr. 2000.

[9] R. Cole and J. Rosenbluth. Voice over IP performance monitoring. *ACM Computer Communication Review*, 31(2), Apr. 2001.

[10] C. Filsfils. Sub-100ms convergence MPLS FRR. RIPE41 meeting, Jan. 2002.

[11] C. Fraleigh, S. Moon, C. Diot, B. Lyles, and F. Tobagi. Packet-level traffic measurements from a tier-1 IP backbone. Technical Report TR01-ATL-110101, Sprint ATL, Nov. 2001.

[12] O. Hagsand, K. Hanson, and I. Marsh. Measuring Internet Telephony quality: where are we today? In *Proc. Globecom'99*, pages 1843–1842, Dec. 1999.

[13] W. Jiang and H. Schulzrinne. Qos measurement of internet real-time multimedia services. Technical Report CUCS-015-99, Columbia University, December 1999.

[14] W. Jiang and H. Schulzrinne. Modeling of packet loss and delay and their effect on real-time multimedia service quality. In *Proceedings of ACM NOSSDAV*, June 2000.

[15] S. Keshav. *An Engineering Approach to Computer Networking*. Addison-Wesley Publishing Company, 1997.

[16] T. J. Kostas, M. S. Borella, I. Sidhu, G. M. Schuster, J. Grabiec, and J. Mahler. Real-time voice over packet-switched networks. *IEEE Network*, pages 18–27, January/February 1998.

[17] A. Markopoulou, F. Tobagi, and M. Karam. Assessment of VoIP quality over internet backbones. In *to appear in Proc. Infocom'02*, 2002.

[18] J. Micheel, S. Donnelly, and I. Graham. Precision timestamping of network packets. In *Proceedings of ACM Sigcomm Internet Measurement Workshop*, Nov. 2001.

[19] R. Mortier. Python routeing toolkit, 2002. http://www.sprintlabs.com/Department/IP-Interworking/Routing.

[20] D. Oran. OSI IS-IS intra-domain routing protocol. RFC 1142, Feb. 1990.

[21] D. Papagiannaki, S. Moon, C. Fraleigh, P. Thiran, F. Tobagi, and C. Diot. Analysis of measured single-hop delay from an operational backbone network. In *to appear in Proc. Infocom'02*, 2002.

[22] V. Paxson. End-to-end routing behavior in the internet. *IEEE/ACM Transactions on Networking*, 5(5):601–615, Oct. 1997.

[23] E. C. Rosen, A. Viswanathan, and R. Callon. Multiprotocol Label Switching Architecture. RFC 3031, Jan. 2001.

[24] M. Yajnik, S. Moon, J. Kurose, and D. Towsley. Measurement and modelling of temporal dependence in packet loss. In *Proc. IEEE Infocom'99*, March 1999. New York, NY.