Can You Hear Me Now?! It Must Be BGP

Nate Kushman nkushman@mit.edu
Srikanth Kandula kandula@mit.edu
Dina Katabi dk@mit.edu

This paper uses empirical results to argue that VoIP's performance is unacceptable and BGP is the main culprit.

1 Introduction

The VoIP market is growing rapidly. Analysts estimate that 20% of voice traffic in North America will run over IP by 2010 [16]. Many businesses are adopting VoIP as it is a cheap substitute for land-line service [17], particularly for long distance calls. Those who do not move to VoIP state quality concerns as a major reason [17], highlighting a need to better understand VoIP performance.

This paper argues that BGP hinders VoIP performance. Our position stems from an empirical study of VoIP quality in which we ask a basic question: how does end-to-end VoIP perform over the Internet? To realistically evaluate the quality of actual VoIP calls over the wide-area Internet, we employ an open source implementation of the popular Google Talk client and run more than 50,000 automated phone calls over the RON testbed [6]. We evaluate the user perceived quality of the actual voice streams using an industry standard signal processing technique called PESQ [20], that approximates the mean opinion score (MOS) of human judges. We then use BGP feeds matching the studied routes to discover whether poor call quality correlates with BGP updates.

Our study reveals the following:

(1) VoIP quality is unacceptable: VoIP is unusable for an VoIP is unintelligible for an average of 15 minutes a day. In contrast to the public phone network, which has a reliability between 99.94% and 99.999% [32], VoIP provides usable quality only about 99% of the time. Further, there is an 8% chance that a user making a long VoIP call will hang up within the first hour due to network outages.

(2) BGP is a major cause of poor voice quality:

- As those familiar with BGP would expect, on average, when a BGP update happens, voice quality falls to an unintelligible level. In more than 50% of BGP-correlated outages, the call remains unintelligible for over four minutes, preventing callers from reestablishing the call.[1]
- Surprisingly however, more than 50% of the periods with unintelligible voice quality occur within 15 minutes of a BGP update, and 97% are within 40 minutes of a BGP update. This is striking given that update events on our paths are, on average, about a day apart.

Many non-BGP causes for poor VoIP performance such as congestion and susceptibility to intra-domain link failures can be alleviated with existing basic QoS and intra-domain recovery techniques (such as MPLS fast reroute) [31]. Without fixing BGP, however, we leave the ISPs with a VoIP service that is usable most of the time, but suffers from frequent outages due to BGP updates. It is unlikely users will replace their land-line service with less expensive VoIP if it is unreliable. This is particularly true for the North American market [17], and is more apparent in the case of business calls, where a dropped call can mean a lost customer. Thus, unless BGP problems are addressed, many potential clients will not switch to VoIP due to concerns about quality.

How do we address quality degradation due to BGP? Inter-domain overlays can potentially alleviate the problem [42, 7]. However, whether overlays can react soon enough –50% of callers hang up after 12 seconds of outage [19]– and without excessive overhead is an open area of research.[2] More importantly, though, inter-domain overlays are not practical solutions for ISPs [13], and it seems likely that the ISP's will be significant players in the VoIP market [14]. Currently, they either derive substantial revenues from carrying VoIP traffic (e.g., Comcast, AOL), or are looking to move into the VoIP market to make money off the unused capacity in their backbones (e.g., Verizon) [41]. However, ISPs have traditionally been unable to build inter-domain overlays, which requires them to pull together their resources and cooperate tightly, creating an ISP federation of some form [13]. Thus, ISPs have an incentive to look for an alternative solution that directly tackles BGP problems.

We believe this work provides the first evidence of the correlation between VoIP performance and BGP updates. Equally importantly, we identify performance problems correlated with BGP that ISP's actually have an economic incentive to solve. This work also motivates the need for novel solutions. Most prior solutions to BGP problems are reactive; they reduce convergence time and the number of

[1] This does not mean that the average BGP event lasts for 4 minutes. Many BGP events are very short and do not cause outages. The ones that cause outages however tend to last for a few minutes.

[2] Skype's overlay is mainly used to by-pass NATs, as opposed to performance problems [7].

messages exchanged during convergence [38, 34, 10, 33]. BGP convergence, however, is limited by various timers (e.g., the MRAI timer is usually set to 30s [36, 15]). Thus, reactive solutions are unlikely to satisfy VoIP's strict real-time constraints, where 50% of callers hang up within 12 seconds of poor voice quality [19]. It seems that proactive BGP solutions such as precomputed fail-over paths [24, 8] and inter-domain multipath routing [46] would be better suited. Future research should explore how to employ these techniques to address current VoIP problems.

2 Related Work

VoIP has recently become an active research area, with existing work in three general categories: VoIP performance studies, studies of Skype's network, and other VoIP work. We believe that this submission is the first paper to correlate VoIP quality with BGP updates. It further differs from prior work in one or more of the following: (1) It employs commercial VoIP application. (2) It measures the user perceived quality of voice, as opposed to network-related metrics like delay and jitter. (3) It is significantly larger than previous studies, with 50,000 calls between 24 locations, in US and Europe.

The prior work closest to ours measures VoIP quality. Earlier work looks only at VoIP performance in a single backbone network. Boutremans et al. [9] examine VoIP's performance on the Sprint network. They find that due to over provisioning, congestion related packet-loss and latency are not significant issues for VoIP performance on a Tier-1 network. Markopoulou et al. [28] use data from 200 phone calls to study VoIP performance across various individual ISPs, and find that about 5% of their calls have inadequate quality. More recent work measures cross-domain performance, but focuses only at packet level statistics to estimate user perception [4, 35, 39, 21, 29]. Some prior work also studies the performance of VoIP overlays. A few papers [26, 5, 35, 39, 22] show that, when possible, using multiple paths or dynamically switching paths based on real-time metrics improves VoIP performance.

Skype's popularity has motivated studies to understand its peer-to-peer VoIP network. Some of these have measured characteristics of current Skype network [7], quantified user satisfaction by looking at call duration [11], or looked at the performance of Skype vs. other VoIP clients [12].

A few papers have examined the performance of VoIP over wireless networks [23], its interaction with TCP and UDP [47], and how to provide the equivalent of "911" with a VoIP service [30].

Finally, much work has shown that both BGP update events occurring in controlled fashion [27] and updates occurring in the wild [25] are strongly correlated with periods of high packet loss, increased path delay and loopy paths [44, 18, 43]. Our study is the first to show that for realtime applications such as VoIP, BGP update are correlated with a *significant* fraction of poor quality periods.

Name/Location	**Upstream Provider(s)**
Amsterdam, Netherlands	Global Crossing (GBLX)
Salt Lake City, Utah	xmission.net → Level3
Chicago, IL	GBLX
CMU, Pittsburgh, PA	AT&T, GBLX, Abilene . . .
Laurel, MD	coloco.com → Cogent
Toronto, Canada	convoke.net → Cogent
Cornell, Ithaca, NY	Abilene . . .
Delta, Canada	bigpipeinc.com → Level3
San Luis Obispo, CA	dwni.net → Cogent
Gatech, Atlanta, GA	Abilene . . .
Austin, TX	hostway.com → Cogent
New York, NY	GBLX
London, UK	GBLX
MIT, Boston, MA	Genuity, Abilene, Cogent, Comcast
San Jose, CA	megapath.net
Mount Vernon, IL	mvn.net
Hillsborough, NC	rr.com
NYU, New York, NY	Abilene . . .
Tacoma, WA	opticfusion.net
New York, NY	speakeasy.net
UCSD, San Deigo, CA	CENIC, Abilene . . .
UMich, Ann Arbor, MI	Abilene . . .
Univ. Utah, Salt Lake City, UT	Abilene . . .
New York, NY	webair.net → GBLX

Table 1—We ran VoIP calls between 24 hosts. The five hosts in bold have BGP feeds and serve as the VoIP servers. Each VoIP call had at least one end-point with a BGP feed

3 VoIP

In this section, we describe the components that VoIP software typically contains, the various algorithms for encoding and decoding voice traffic and the industry standard method to measure voice call quality.

3.1 VoIP Software

Fig. 1 shows the four layers that make up a typical VoIP connection. At the lowest layer are mechanisms such as STUN or ICE that punch holes through NATs/firewalls. Above this is a signaling protocol, typically SIP or H.323, that creates the RTP (Real Time Protocol)[37] connection, negotiates the CODEC to be used, and attempts to keep the sender and receiver synchronized using periodic RTCP messages. Just above the signaling protocol, is a library such as Linphone [1] that handles the real-time scheduling– typically, once every 20ms this library picks a sample from the sound input, encodes it using the pre-negotiated CODEC, and hands it off to be sent over UDP. At the receiving end, packets arriving off the wire are re-assembled in a jitter-concealing buffer, decoded and written to the sound card, again once every 20ms. Finally, clients such as Skype and Google Talk handle all user interaction, peer-to-peer discovery of the caller/callee, and optionally route the call over an overlay.

Most popular end-to-end VoIP clients are closed source, including Skype, MSN Messenger, Yahoo Messenger, AIM and Google Talk. Some, such as Skype, even implement proprietary protocols for all layers. Google, however, provides Libjingle, an open source "example" implementation of the Google Talk signaling and NAT hole-

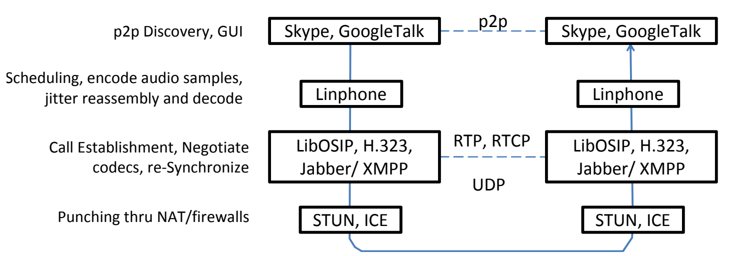

Figure 1—Components in a typical VoIP use scenario. For our experiments we used a customized Linphone client, running on top of libOSIP.

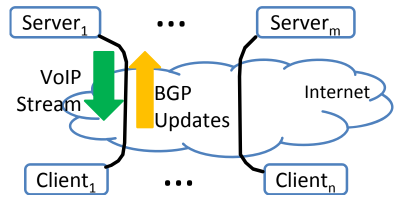

Figure 2—We pick nodes with BGP feeds to be the servers of VoIP streams. Recall that BGP updates flow in the reverse direction of the data path. This ensures that we collect BGP updates for the paths taken by the VoIP streams.

punching protocol, which uses the open source LinPhone project [1] as its real-time component. Since all of our testbed machines have their own public IP addresses, we did not need the p2p discovery and NAT hole-punching functionality. Thus, in our experiments we used a modified Linphone client/server which uses the SIP signaling protocol. Our modified receivers dump the received VoIP stream to file instead of writing to the audio device.

3.2 CODECs

The choice of a VoIP CODEC trades off bandwidth for voice quality. Recent CODECs claim to provide reasonable quality, while using as little as 6kbps. Table 2 summarizes the available options. Popular VoIP clients often support several of these CODECs, and sometimes adaptively pick a CODEC based on path quality. Our experiments use the G.711 (PCMU) CODEC, which is the highest quality CODEC in common use and uses about 64Kbps of bandwidth. We choose this CODEC because it is the default choice for most VoIP software packages and is the CODEC used by the PSTN. Additionally, the PCMU CODEC specification advocates filling in lost frames (packets) with the previous frame. Linphone did not implement this feature, so we added it after ensuring that it does improve quality.

3.3 Measuring VoIP call quality

The Mean Opinion Score (MOS) is the industry standard for measuring call quality [2]. It is expressed on a scale of 1 to 5, as follows.

MOS	Rating	Perceived Quality
4-5	Excellent	Toll Quality
3-4	Good	Cell Phone Quality
< 3	Fair	Unacceptable
< 2	Bad	Unintelligible

Ideally, the MOS is computed as the mean of the scores given by human judges who rate VoIP samples on a scale from 1 (bad) to 5 (excellent). However, human studies need controlled environments and are quite expensive. Instead, end-to-end tests of voice quality use standardized automated techniques[45], the most advanced of which is a signal processing based algorithm called PESQ (Perceptual Evaluation of Sound Quality) [20, 40]. PESQ compares the voice sample received over VoIP with the "original" voice

CODEC	Data rate	MOS	Notes
G.711	64 kbit/s	4.1	Toll Quality, Small delay in coding
G.729	8 kbit/s	$\approx$3.9	$\approx$ 10ms coding delay
G.723	5.3/6.3 kbit/s	$\approx$3.6-3.8	FEC + About 30ms coding delay
Speex	2.2-44.2 kbit/s	$\approx$3.8	FEC + Variable Bit Rate for Low Bandwidth
iLBC	13.3/15.2 kbit/s	$\approx$3.8	Upon packet loss, add FEC and increase rate

Table 2—Some popular VoIP codecs. VoIP clients typically support many of these CODECs and some adaptively switch codecs.

sample to yield a MOS that has been shown to be highly correlated with MOS scores from human judges[20]. We use the PESQ algorithm for our experiments. When the call is dropped, or the client cannot reconnect, there is no received voice stream to compute the PESQ score. We use a MOS of 0 for all such samples.

4 Experimental Setup

To measure the effect of naturally occurring BGP updates on VoIP, in as realistic a way as possible, we ran over 50,000 one hour long phone calls along approximately 100 paths between RON hosts for three months. We also collect the BGP updates affecting these paths, and correlate the performance of the VoIP streams with BGP updates.

(a) Measured Paths: We used the measurement sites from the RON [6] testbed shown in Table 1. We chose RON because BGP updates are available at several nodes. Also, the relatively light load on RON nodes allows the real-time VoIP clients to function with fewer CPU starvation incidents. To avoid bias, we only include paths that do not have both ends connected to Internet2 [3].

(b) Call Configuration: We run a VoIP server on each of the five nodes with BGP feeds. The server is contacted by clients running on other nodes. We establish one SIP session over each of the paths available. For the practical reason of avoiding extremely long-running applications, we re-establish VoIP sessions every hour.

(c) Experiment Setup: Fig. 2 shows our experimental setup. Since the BGP feeds tell us only about routing changes for packets leaving the machine and give us no information about the routing of packets coming to the machine, we chose to make one sided phone calls with all voice data traveling from the machines with BGP feeds. This means we do not capture the effect of BGP updates on the reverse path. Though the RTP protocol used by our VoIP client sends control packets in the reverse direction,

Term	Definition
Sample	3.5 second long received voice
Outage	contiguous period of samples with MOS < 2
Call Abandonment	premature hang ups caused by outages

Table 3—Definitions of terms used in the paper.

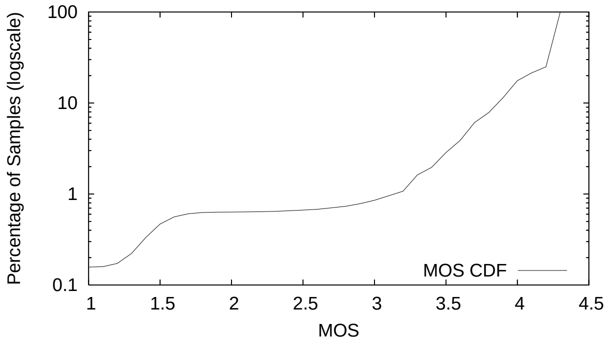

Figure 3—CDF of the MOS over all samples in one day, shown with a log scale. It reveals that nearly 1% of samples have a MOS below 2, which reflects either an outage or unintelligible voice.

such packets are infrequent and the protocol is more resilient to losses and/or high latency in this direction [37]. So we believe that most routing problems impacting the connection occur on the forward path.

(d) Choosing Voice to Transmit: We pick a standard 3.5 second spoken sound sample of an English sentence (one utterance) provided with the ITU reference PESQ implementation. To simulate a voice call, we modified the server to replay this raw audio file repeatedly.

(e) Eliminating CPU Scheduling Issues: We have observed many instances in which CPU scheduling at the end-hosts introduces excessive jitter. For example, the server sending the packets is not scheduled for 40ms, or the client receiving the packets is not scheduled for 60ms. Since our focus here is on the impact of network events on VoIP performance, we have discarded samples that seem to have been affected by CPU scheduling issues.

5 RESULTS

Is VoIP hindered by BGP? We answer this question by examining whether VoIP performs well, and if not, then to what extent the problems are correlated with BGP updates. Our study supports two main conclusions. First, it shows that VoIP on the current Internet though usable, does not perform well enough to replace the Public Switched Telephone Network (PSTN). Second, most VoIP performance problems are correlated with BGP updates.

Our description of the results uses the terms: “sample”, “outage”, and “call abandonment”, defined in Table 3.

5.1 Call Quality & Availability

Our results show that voice quality in VoIP is relatively good, but availability is significantly lower than the PSTN. Figure 3 shows that 99% of call samples have a quality similar or better than cell phones –i.e., a MOS higher than 3. On the flip side, VoIP availability is low. Figure 3 shows that about 1% of samples have a MOS lower than 2, which implies an unintelligible utterance, and hence an outage. This means that users experience on average 10-15 minutes a day of outage. Though this level of availability is acceptable for a free casual-use service, it is not reliable enough to replace PSTN land-line phones, which are available for 99.94% to 99.999% of the time [32].

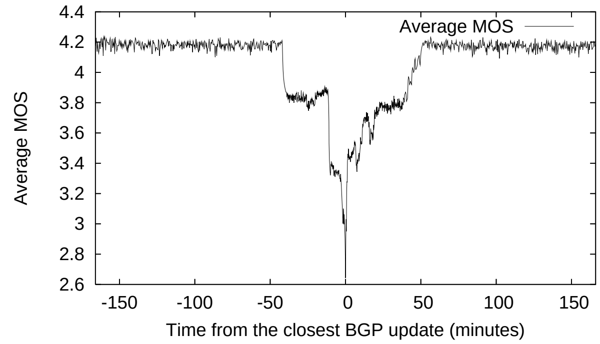

Figure 4—Average MOS across all samples as a function of the time between the sample and the closest BGP update. The figure shows that near a BGP update, the avg. MOS drops significantly. Both outages (MOS < 2) and bad quality samples (MOS < 3) occur around BGP updates.

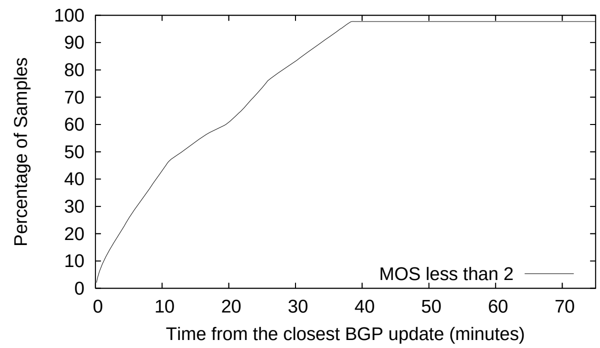

Figure 5—CDF of the time from closest BGP update (in minutes), taken over all outage samples. It shows that more than 50% of outage samples are within 15 minutes of a BGP update, and almost all of them are within 40 minutes of a BGP update. Note BGP updates on our paths are spaced by about a day.

5.1.1 *Correlation with BGP updates*

To show that drops in MOS are highly correlated with BGP updates, we must show two things. First, we must show that, when a BGP update occurs, on average there is a significant drop in MOS. Additionally, we must show the opposite, that when MOS score drops, it is likely that a BGP update is nearby.

(1) BGP Update $\Rightarrow$ Low MOS: Figure 4 shows that the average MOS near a BGP update drops from over 4.2 for samples far from a BGP update, to less than 3 for samples near a BGP update. Thus for samples far from a BGP update, the average MOS is toll quality, while near a BGP update the average MOS drops below cell phone quality.

(2) Low MOS $\Rightarrow$ Nearby BGP Update: Figure 5 shows

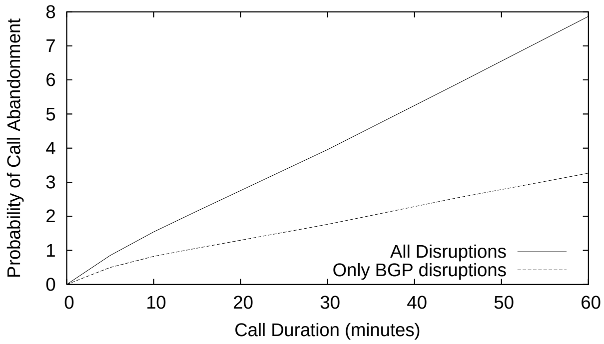

Figure 6—Percentage of calls abandoned as a function of call length for all outages, and only outages correlated with BGP updates. Percentages are computed using Eq. 2 and averaged across all possible call start times on all paths. The figure shows that 8% of one-hour calls are abandoned because of poor quality.

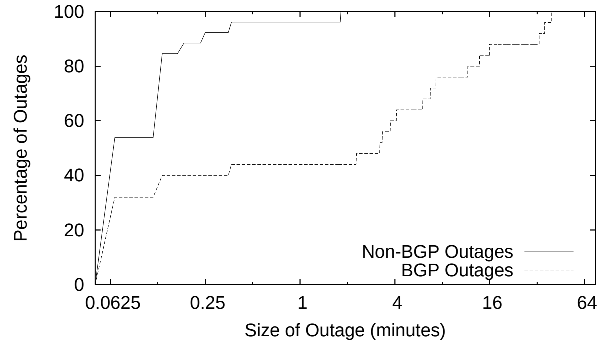

Figure 7—CDF of outage length (in minutes) for samples correlated and uncorrelated with BGP updates. It shows that outages correlated with BGP last longer than those uncorrelated with BGP.

a surprising result; more than 50% of the periods with outages (i.e., a MOS < 2) occur within 15 minutes of a BGP update, and 97% are within 40 minutes of a BGP update. This is striking given that update events on our paths are, on average, spaced about a day apart.

To summarize, VoIP quality is good but its availability is bad. The vast majority of unavailability samples are correlated with BGP updates. More specifically, when a link goes down or comes up, BGP explores alternative paths. Before converging to a stable path, BGP may incur periods of drops, delay, loops, and transient disconnectivity. These dynamics are behind most of VoIP outages.[3]

5.2 Call Abandonment

Another important metric for VoIP performance is the ability to conduct long conversations without being forced to abandon the call. Poor VoIP quality leads callers to prematurely hang up. Since we do not have actual users, we could not measure hang up rates. Instead, we analyze this effect using the telephony industry's standard for estimating hang up rates. ITU standard E.855 [19] estimates that the likelihood a user does not hang up the call is exponentially distributed with the length of outage, d, in seconds:

$$P[\text{hold on}] = e^{-d/17.26}. \quad (1)$$

We combine this estimate with our voice data to estimate the likelihood that a user can conduct a conversation of duration x without hanging up. A call of duration x beginning at time t will not be abandoned only if the user does not hang-up during every one of the outages that happen between t and $t+x$. Let $d_1, d_2, \ldots$ be the outages that occur between t and $t+x$, then the probability of not hanging up during that period $P_x(t)$ is:

$$P_x(t) = \prod_i P(\text{hold on during } d_i) = e^{-\frac{\sum_i d_i}{17.26}}. \quad (2)$$

[3]We acknowledge that correlation does not necessarily imply causality. There is no way to check causality without active measurements (i.e., without injecting BGP updates). But given how BGP and the Internet work, it is unlikely to see such high correlation without causality.

We then average over many starting times t to estimate P_x – the probability that a call of duration x goes through.

Computing the probability of call abandonment as a function of the call duration shows that in 8% of one-hour calls (1 out of 12), the caller will prematurely hang up because of outages. More generally, Figure 6 shows the probability of call abandonment, $1 - P_x$, as a function of call duration, computed using Eq. 2. Given that 97% of outage samples are correlated with BGP, one might expect outages caused by BGP to dominate the number of abandoned calls. But, BGP outages occur far more clustered than non-BGP outages. The exponential nature of Eq. 2 means that a single outage of length 7s is likely to force fewer calls to be abandoned than two outages each of length 3.5s occurring at different times. Nevertheless, about half of call abandonment is correlated with BGP updates.

5.3 Call-Back Success

Looking only at call abandonment under-represents the frustration caused by BGP updates, because outages correlated with BGP updates tend to be far longer than outages not correlated with BGP updates. Specifically, Figure 7 shows that all non-BGP outages last less than two minutes, while about 30% of BGP outages persist for greater than 10 minutes, meaning that during BGP outages, attempts to call back after the initial abandonment will likely fail.

6 Concluding Remarks

We have shown that VoIP outages correlated with BGP updates occur often enough, cause enough calls to be abandoned, and create long enough unavailability that we believe it is difficult to replace PSTN based service with end-to-end VoIP without fixing the performance problems correlated with BGP.

We recognize, however, that our study of VoIP performance is limited in a few ways. First, we stream voice only in one direction. Two-way voice is the norm in telephony and seems to demand more stringent performance – for example users will hang-up a call when a problem happens on either the forward or the reverse paths. Second, we focus on a relatively small number of paths. To obtain results with high statistical confidence, we need to look at many more than a hundred paths for much longer than three months. Finally, newer CODECs such as iLBC and Speex, employ sophisticated forward error correction techniques that can recover from random packet losses – this has the potential to improve overall performance of VoIP but may also make it more crucial to avoid the bursty losses caused by BGP events.

Nonetheless, our results motivate a change in perspective on both the importance of, and the method for, solving the performance problems associated with BGP updates. The ISPs' desire to capture the VoIP market may finally provide the necessary incentive for them to fix BGP. However, we believe that reactive solutions, which attempt to reduce convergence time or the number of messages exchanged during path exploration, cannot operate on the short time scale demanded by VoIP. User studies show that 50% of people will hang up a phone call after less than 12 seconds of poor voice quality [19], yet BGP convergence is limited by timers such as the MRAI timer that defaults to a 30 second timeout value [36]. This leads us to believe that methods allowing proactive discovery of fail-over paths will be required.

REFERENCES

[1] Linphone: Telephony on Linux. http://www.linphone.org.

[2] Mean Opinion Score. http://en.wikipedia.org/wiki/Mean_Opinion_Score.

[3] Abilene. http://monon.uits.iupui.edu/.

[4] Y. Amir, C. Danilov, S. Goose, D. Hedqvist, and A. Terzis. An Overlay Architecture for High Quality VoIP Streams.

[5] Y. Amir et al. 1-800-OVERLAYS: Using Overlay Networks to Improve VoIP Quality.

[6] D. Andersen, H. Balakrishnan, M. F. Kaashoek, and R. Morris. Resilient Overlay Networks. In *SOSP*, 2001.

[7] S. Baset and H. Schulzrinne. Analysis of the Skype Peer to Peer Internet Telephony Protocol. In *INFOCOM*, 2006.

[8] O. Bonaventure, C. Filsfils, and P. Francois. Achieving sub-50ms Recovery upon BGP Peering Link Failures. In *Co-Next*, 2005.

[9] C. Boutremans, G. Iannaccone, and C. Diot. Impact of link failures on VoIP performance. In *NOSSDAV*, 2002.

[10] A. Bremler-Barr, Y. Afek, and S. Schwarz. Improved BGP Convergence via Ghost Flushing. In *INFOCOM*, 2003.

[11] K.-T. Chen et al. Quantifying Skype User Satisfaction. In *SIGCOMM*, August 2006.

[12] W.-H. Chiang et al. A performance study of voip applications: Msn vs. skype. In *MULTICOMM*, 2006.

[13] D. Clark, W. Lehr, P. Faratin, S. Bauer, and J. Wroclawski. The Growth of Internet Overlay Networks: Implications for Architecture, Industry Structure and Policy. In *TPRC*, 2005.

[14] T. Espiner. Skypes's market share halves, February 2006.

[15] A. Feldmann, H. Kong, O. Maennel, and A. Tudor. Measuring bgp pass-through times. In *Proc. of the Passive and Active Measurement Workshop (PAM)*, 2004.

[16] Frost and Sullivan. North American Wholesale Long Distance Voice Service Market - Gaining a Competitive Advantage Through Migration to VoIP, Dec 2004.

[17] Frost and Sullivan. Trends in Wireline Substitution - North American Markets, August 2005.

[18] P. Huang, A. Feldmann, and W. Willinger. A non-intrusive, wavelet-based approach to detecting network performance problems. In *Proc. of ACM SIGCOMM Internet Measurement Workshop 2001, San Francisco Bay Area*, Nov. 2001.

[19] International Telecommunication Union. Connection Integrity Objective for International Telephone Service. E.855., 1988.

[20] International Telecommunication Union. Perceptual Evaluation of Speech Quality. In *ITU-T Rec. P.862*, 2001.

[21] W. Jiang and H. Schulzrinne. Assessment of VoIP service availability in the current Internet. In *PAM*, 2003.

[22] M. Karol, P. Krishnan, and J. Li. Using Overlay Networks to Improve VoIP Reliability. In *LNCS*, 2004.

[23] T. Kawata, S. Shin, A. G. Forte, and H. Schulzrinne. Improving the Capacity for VoIP Traffic in IEEE 802.11 Networks with Dynamic PCF. In *IEEE WCNC*, 2005.

[24] N. Kushman, S. Kandula, D. Katabi, B. Maggs, and J. Wroclawski. RT-BGP: Supporting Inter-domain Real-Time Applications. In *Poster NSDI*, 2006.

[25] C. Labovitz, A. Ahuja, A. Bose, and F. Jahanian. Delayed internet routing convergence. In *SIGCOMM*, pages 175–187, 2000.

[26] Y. Liang, E. Steinbach, and B. Girod. Multi-stream voice over IP using Packet Path Diversity. In *IEEE Multimedia Signal Processing*, 2001.

[27] Z. Mao, R. Govindan, G. Varghese, and R. Katz. Route flap damping exacerbates internet routing convergence. In *ACM SIGCOMM*, 2002.

[28] A. Markopoulou, F. Tobagi, and M. Karam. Assessment of VoIP Quality over Internet Backbones. In *INFOCOM*, 2002.

[29] I. Marsh and F. Li. Wide Area Measurements of Voice over IP Quality. Technical report, Swedish Inst of Computer Science, 2003.

[30] M. Mintz-Habib, A. S. Rawat, H. Schulzrinne, and X. Wu. A VoIP Emergency Services Architecture and Prototype. In *ICCCN*, 2005.

[31] E. Osborne and A. Simha. *Traffic Engineering with MPLS*. Cisco Press, 2002.

[32] PacketCable. VoIP Availability and Reliability Model for PacketCable Architecture. Technical Report PKT-TR-VoIPPAR-W05-001012, 2000.

[33] D. Pei et al. Improving BGP Convergence Through Consistency Assertions. In *INFOCOM*, 2002.

[34] D. Pei et al. BGP-RCN: Improving BGP Convergence through Root Cause Notification. Technical Report TR-030047, UCLA, 2003.

[35] R. K. Rajendran et al. Performance Optimization of VoIP using an Overlay Network. Technical report, NEC, 2005.

[36] Y. Rekhter and T. Li. Border Gateway Protocol 4. In *RFC 1771*, 1995.

[37] H. Schulzrinne et al. RTP: A Transport Protocol for Real-Time Applications. In *IETF RFC 1889*, 1996.

[38] L. Subramanian, M. Caesar, C. T. Ee, M. Handley, M. Mao, S. Shenker, and I. Stoica. HLP: A Next-generation Interdomain Routing Protocol. In *SIGCOMM*, 2005.

[39] S. Tao et al. Improving VoIP Quality Through Path Switching. In *INFOCOM*, 2005.

[40] M. Varela, I. Marsh, and B. Grnvall. A Systematic Study of PESQ's Performance. In *MESAQIN*, 2006.

[41] Verizon Rings in Next Generation of Voice Service. http://newscenter.verizon.com/proactive/newsroom/release.vtml?id=86115.

[42] Vonage – IP Telephony/Voice over IP. http://www.cisco.com/en/US/tech/tk652/tk701/technologies_case_study09186a00800b559e.shtml.

[43] F. Wang, N. Feamster, and L. Gao. Quantifying the effects of routing dynamics on end-to-end internet path failures. In *UMASS Technical report (TR-05-CSE-03)*, 2006.

[44] F. Wang, Z. M. Mao, J. W. L. Gao, and R. Bush. A Measurement Study on the Impact of Routing Events on End-to-End Internet Path Performance. In *SIGCOMM*, 2006.

[45] A. Wittman. Assuring VoIP Quality: Not There Yet. IT Architect, March 2005.

[46] W. Xu and J. Rexford. Multi-path Interdomain Routing. In *SIGCOMM*, 2006.

[47] X. Zhang and H. Schulzrinne. Voice over TCP and UDP. Technical Report CUCS-033-04, Columbia Univ., 2004.