1-800-CALL-H.E.P.

Experiences on a Voice-over-IP Test Bed.

W. Matthews[1], *L. Cottrell*[1], *R. Nitzan*[2]

[1] Stanford Linear Accelerator Center
[2] Energy Sciences Network

Abstract

Highly interactive Internet applications such as Voice-over-IP are extremely sensitive to network performance. Even on high performance research networks, many cases will require the use of differentiated services to achieve high (toll) quality conversations.

In this talk we will describe a test bed over the Energy Sciences network (ESnet) between Lawrence Berkeley National Laboratory (LBNL), Stanford Linear Accelerator Center (SLAC), Argonne National Laboratory (ANL) and Sandia National Laboratory (Sandia).

In particular the characteristics of Voice-over-IP calls between LBNL and SLAC will be reviewed and the effect of low, moderate and high congestion on the link will be quantified. The use of Per Hop Behavior (PHB) in IP headers with Weighted Fair Queuing (WFQ) in routers and the benefit they provide will be explained.

A model of flows and performance will be presented and new techniques to predict the quality of calls are under development and will be reviewed. Comparisons with telephone reliability will be discussed and the feasability of wide spread deployment of VoIP in HEP will be considered.

Keywords: Network, Performance, Quality of Service

1 Introduction

The potential cost savings of free telephone calls for laboratories and universities throughout the world is reason enough to trial a VoIP test bed. In addition the evolution of the Internet and in particular the requirements of demanding applications force the community to explore Quality of Service (QoS) techniques such as differentiated services and dedicated bandwidth.

2 The VoIP Test bed

Lawrence Berkeley National Laboratory (LBNL), Stanford Linear Accelerator Center (SLAC), Argonne National Laboratory (ANL) and Sandia National Laboratory (Sandia) are participating in a VoIP test bed carved from the ESnet ATM backbone. 3.5 Mbps is allocated to the test bed.

At each site, a regular telephone connected to a PBX is used with an access code. The PBX is connected to a router and once the access code is provided, the connection is switched to the test bed rather than the usual telephone network. Policing is done at the site edge router and Committed Access Rate (CAR) is applied to the VoIP packets. Weighted Fair Queuing (WFQ) in a router at the edge of the ATM cloud reads the CAR setting and gives priority to the VoIP traffic.

3 A Typical VoIP Call

The router connected to the PBX at the sender site transmits TCP packets to the router connected to the PBX at the destination site and waits for acknowledgment that the remote port is open. This acknowledgment signifies a connection is made and the call has been initiated. Once the call is established, the voice of the caller is encoded into packets using ITU G.729 and sent over UDP.

The results presented and discussed in this paper were conducted between SLAC and LBNL, mostly calling recorded messages.

4 Performance on the VoIP Test Bed

An understanding of the underlying network performance can be extrapolated to build a picture of the performance of any application[1]. Therefore the methodology is to measure packet loss and latency caused by congestion at the bottleneck on the testbed and relate this to the user's perception of the quality and usability of a call. Hence, a quantitative measurement of the quality of the call can be derived, and a set of conditions for network performance required to provide high quality application performance can be defined.

The aim of this work is to understand under what conditions, ie the range of packet loss and RTT that is acceptable to VoIP.

Packet Loss is important because even moderate packet loss during a communication will make the conversation patchy and hard to follow. The `PingER` project defines packet loss less than 1% to be good and suitable for highly interactive applications. Packet loss induced by high Best-Effort (B.E) traffic, ie the "traditional" Internet, is almost entirely the effect of queuing in routers. Packet loss is detected by looking for gaps in the RTP sequence number sniffed by using the `sniffit` program.

Delay and round trip time are important because the end-users must be able to synchronise their conversation. Silence can be mean one person has finished speaking and is waiting for a reply, or can be caused by too much delay and the speakers begin to talk at the same time. A study by the International Telecommunications Union (ITU) defines 300ms [1] to be the required round trip time for good voice communication. Delay is the effect of the speed of light in glass and queuing in routers. Also a significant extra delay was observed when congestion arose due to a bug in the cisco code, but this has been fixed. Two way delay is measured by the `ping` round trip time (RTT).

Audio such as the VoIP application is particularly sensitive to the variation in packet arrival. Inter-Packet-Delay-Variation (IPDV) [2] is used, and jitter is defined to be the magnitude of the IPDV. Jitter is measured by counting the time of the incoming packet with the `TCPDUMP` program. On an unloaded link, the jitter has a sharp peak around 19.9ms. The Inter-quartile-range (IQR) of the jitter or the IPDV can be a simple but robust measure of the quality of the call.

On the unloaded testbed packet loss is negligible, RTT and are very low. Congestion, ie sufficient traffic to cause packet loss and delay above what would be present without the traffic, is simulated by injecting B.E. traffic onto the testbed using a modified version of `TTCP` or the Netcom System's Smart Bit Device.

When CAR is enabled, the call remains clear as the B.E. traffic begins to congest the line and B.E. packets are dropped. Even with very high B.E. packet loss the call remains high quality.

It would take 56 concurrent calls to fill the test bed, hence studies of the performance due to high levels of prioritized traffic have not yet been assessed. However, even prioritized packets are not guarenteed to arrive, so it is expected that competing traffic, even with CAR enabled, would behave similarly to competing B.E. packets.

Another metric to consider is error-free-seconds. This is a measurement often used by telephone companies. Studies on the testbed to gauge the effect of load on error-free-seconds has

[1]but rarely does the performance of a specific application give detailed knowledge of the packet-level network performance

not been conducted, but data from the Surveyor project was used to determine performance across ESnet.

5 Real World Performance

Long term studies from the PingER project shows packet loss continues to decline and Round Trip time continually improves as upgrades in the Internet Infrastructure stays one step ahead of the explosion in traffic. The metrics discussed in section 4 allow us to determine the links where VoIP could be succesfully used. For most labs, the RTT is below the ITU [1] threshold and therefore there are many candidates for suitable for VoIP. However packet loss between networks is probably too high and there is likely too much jitter.

The median error free seconds between SLAC, FNAL. CMU and CERN is 99.89%. This compares favourably with the high 99% error free seconds quoted by most telephone companies.

6 Differentiated Services

Packet Loss, and the consequent loss of voice is, of course, unnacceptable during a VoIP call. Some kind of quality of service is essential. The test bed uses Weighted Fair Queueing and Committed Access Rate.

Per-Hop Behaviour and Expedited Forwarding (E.F.) requires an extensive Infrastructure including a Bandwidth Broker.

High performance links such as the ESnet backbone are often far from congested and often have significant head room, so it may seem QoS is not needed. However, light packet loss is not unheard of and on occassion packet loss can rise to a level that would seriously effect highly interactive sessions. Also router algorithms such as random early drop (RED) has an unknown effect on UDP traffic. Many ISP's have RED switched on, but ESnet doesn't.

7 Production Service

There are a number of issues to be addressed before a production VoIP service can be made available for all users. How could/should applications set the CAR or PHB bits ? How will it be policed ? Obviously everyone will wish their packets to get highest priority. The current procedure is bits set by an application are over-written or ignored.

Rather than an on-demand production service, it is more likely the test bed will be extended, perhaps to CNAF/Italy and other sites. Hardware must be installed to provide the services to a site so extending the service to sites who ask for it need not be an issue. Creating a VoIP service across networks will be a great challange, Inter-network differentiated services is particularly immature and it is unlikely Commerical ISP's will not co-operate any time soon.

8 Conclusion

VoIP calls can be high (at least high enough) quality cheap alternative to long-distance and especially international calling. Currently the volume of calls is very low, but if the technology begins to be widely deployed it is expected the telephone companies will react to protect their investments.

9 Further Work

The next steps are to use the test bed to determine the window and how many consecutive VoIP calls can be made. Also the behaviour of CAR at high bandwidth and the effect of RED on UDP must be studied. Extending the testbed and perhaps collaborating with the Qbone project will be explored. Updates will be posted on the project web page at

`http://www-iepm.slac.stanford.edu/monitoring/voip/`

References

1 International Telecommunications Union, "One-Way Transmission Time", ITU-T Recommendation G.114, February 1996.

2 C. Demichelis and P. Chimento, "Instantaneous Packet Delay Variation Metric for IPPM", IETF draft (Work in progress).