Implementation of Quality of Service in VoIP

Indranil Bose
The University of Hong Kong, Hong Kong

Fong Man Chun Thomas
The University of Hong Kong, Hong Kong

INTRODUCTION

Today, Internet technologies have pervaded every corner of our society. More and more people are benefiting from the Internet in one way or the other. One of the current Internet technologies that may benefit us greatly is voice over Internet protocol (VoIP). According to Hardy (2003, p. 2), VoIP is "the interactive voice exchange capability carried over packet-switched transport employing the Internet protocol." With VoIP technology, one can call anyone in this world at a lower cost, compared to traditional telephone systems. However, VoIP technology has one significant drawback. It has a low degree of reliability. From experimental results it is known that VoIP can achieve only 98% reliability. The service down time per year for VoIP is almost 20 working days (175 hours). For most companies and government organizations, such a degree of reliability is unacceptable since the traditional telephone system can achieve 99.999% reliability with a service down time of only five minutes per year (Kos, Klepec, & Tomaxic, 2005). As a result, quality of service (QoS) is an important concept for VoIP. Using QoS, VoIP may be able to overcome its limitation in reliability.

QoS is often defined as the capability to provide resource assurance and service differentiation in a network. The definition includes two important terms—resource assurance and service differentiation. Resource assurance provides a guarantee about the amount of network resources requested by the user. On the other hand, service differentiation provides higher priority of getting network resources to those applications that have critical latency constraints. Given the importance of low latency for voice communication, it is not difficult to predict that QoS will assume greater importance in the VoIP industry as this technology gains popularity in the mass market. It is reported that VoIP is aggressively growing, and this growth is expected to continue in the coming years.

BACKGROUND

Today's Internet provides best-effort services to all its applications and cannot provide any resources guarantee to applications (Kurose & Ross, 2004). Let us discuss an example to illustrate this concept. Imagine that a network uses links with capacity of 2 Mbps and supports two users—John and Peter. John uses a 1.5 Mbps VoIP application and communicates with Peter. Normally, the VoIP application works because the capacity of the link is 2 Mbps (>1.5Mbps). However, when John uses the FTP application at the same time (assuming that the FTP application needs more than 0.5 Mbps), the VoIP application cannot get the amount of resources it needs. This leads to congestion in the network and time delay in voice communication. Therefore, end-to-end QoS is required for providing resource guarantee and service differentiation in order to enhance the reliability of the VoIP system (Fineberg, 2005).

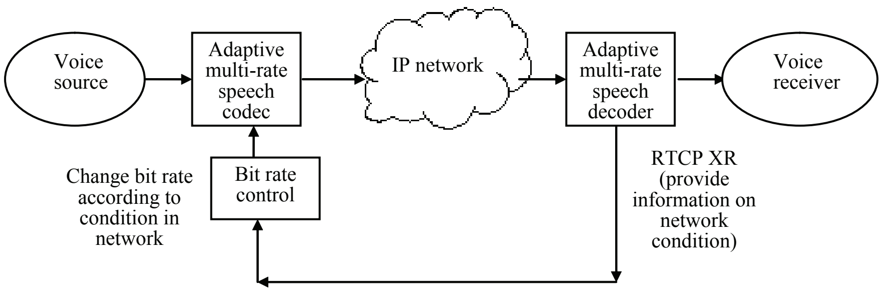
Approaches to provide QoS can be divided into three categories. They are (1) at sender side, (2) inside network, and (3) at receiver side (Wang, 2001). At the sender side, one popular way is by using adaptive multi-rate speech codec. In the network, four core technologies for providing QoS are integrated services (IntServ), differentiated services (DiffServ), multiprotocol label switching (MPLS), and traffic engineering. At the receiver side, the common way for providing QoS is optimization of the design of receiver buffer. The various methods for implementation of QoS for VoIP are described in the next section.

IMPLEMENTATION OF QoS

Methods for QoS Implementation

Presently there are three different approaches for QoS implementation. They can be categorized into QoS at

Figure 1. QoS for VoIP at the sender side



sender side, QoS inside network, and QoS at receiver side.

QoS at Sender Side

At the sender side, QoS can be implemented by using adaptive multi-rate speech codec. A codec is also called coder-decoder, which encodes voice signals to packets or decodes packets to voice signals. When it is an adaptive multi-rate one, it means the codec's encoding bit rate can react with the feedback from congestion in the network (Qiao, Sun, Heilemann, & Ifeachor, 2004). With feedback of congestion the encoding bit rate decreases. Otherwise, the encoding bit rate increases. Figure 1 shows how QoS for VoIP can be provided at the sender side.

In Figure 1 we can see that there is a voice source at the left hand side of the diagram. When the voice source is connected to the adaptive multi-rate speech codec, the codec sends out voice packets to the IP network, and finally voice packets come to the adaptive multi-rate speech decoder. The decoder decodes the voice packets to the voice signal, and finally the receiver retrieves the voice packets. When voice packets travel through the IP network, RTCP XRs (real time control protocol extended reports) are generated in the routers inside the network, and they provide information about network congestion. When the bit rate controller receives the RTCP reports, if the RTCP reports show no/little congestion, the sender side increases its encoding bit rate. Otherwise, the sender decreases its encoding bit rate. This method can provide better resource guarantee because it follows the principle of preemptive congestion avoidance. Some researchers have even proposed the use of a call agent that can selectively compress the voice packets on the sender side based on the status of congestion in the network (Galiotos, Dagiuklas, & Arkadianos, 2002).

QoS Inside Network

Inside the network, there are many approaches for implementing QoS. The most important technologies are integrated services, differentiated services, multi-protocol label switching, and traffic engineering.

IntServ and DiffServ are two architectures that provide resource guarantee and service differentiation. On the other hand, MPLS and traffic engineering are a set of tools for managing the bandwidth and optimizing the performance of the network. The optimization in the design of the receiver buffer involves achieving a good balance between packet loss (due to late packet arrival) and play-out delay. For example, a larger buffer size means less packet loss (thus better quality) but more play-out delay.

Integrated service is a service architecture developed by the Internet Engineering Task Force (IETF) in the early 1990s. It aims at providing QoS for real-time applications such as video conferencing. IntServ is based on per-flow resources reservation. Per-flow resources reservation means that the two end users (sender and receiver) need to make a resource reservation inside the network before using the application. Resource reservation setup protocol (RSVP) is the standard protocol used for resource reservation. IntServ provides two models—guaranteed service model and controlled load service model. Guaranteed service model provides worst-case delay bound for applications. It is designed for applications that need stringent time constraints. Controlled load service model provides less firm delay bounds to applications, and it is similar to but better than the current best effort model.

Related Content

Multi-Sensor Motion Fusion Using Deep Neural Network Learning
Xinyao Sun, Anup Basuand Irene Cheng (2017). *International Journal of Multimedia Data Engineering and Management (pp. 1-18).*
www.irma-international.org/article/multi-sensor-motion-fusion-using-deep-neural-network-learning/187137

The Research on Shape Context Based on Gait Sequence Image
Rong Wang, Yongkang Liuand Mengnan Hu (2018). *International Journal of Multimedia Data Engineering and Management (pp. 21-35).*
www.irma-international.org/article/the-research-on-shape-context-based-on-gait-sequence-image/201914

Automatic Pitch Type Recognition System from Single-View Video Sequences of Baseball Broadcast Videos
Masaki Takahashi, Mahito Fujii, Masahiro Shibata, Nobuyuki Yagiand Shin'ichi Satoh (2012). *Methods and Innovations for Multimedia Database Content Management (pp. 119-142).*
www.irma-international.org/chapter/automatic-pitch-type-recognition-system/66691

Brain Neuron Network Extraction and Analysis of Live Mice from Imaging Videos
Ruichi Yu, Jui-Hsin (Larry) Lai, Shun-Xuan Wangand Ching-Yung Lin (2017). *International Journal of Multimedia Data Engineering and Management (pp. 1-20).*
www.irma-international.org/article/brain-neuron-network-extraction-and-analysis-of-live-mice-from-imaging-videos/182648

Adaptive Retransmission Scheme for Video Streaming Applications
Árpád Huszák (2009). *Handbook of Research on Mobile Multimedia, Second Edition (pp. 248-262).*
www.irma-international.org/chapter/adaptive-retransmission-scheme-video-streaming/21008