Voice Communication Across the Internet: A Network Voice Terminal

H. Schulzrinne

CS Technical Report 92-50
July 1992

Voice Communication Across the Internet: A Network Voice Terminal*

Henning Schulzrinne
Department of Electrical and Computer Engineering
Department of Computer Science
University of Massachusetts
Amherst, MA 01003
hgschulz@cs.umass.edu

July 29, 1992

Abstract

Voice conferencing has attracted interest as a useful and viable first real-time application on the Internet. This report describes NEVOT a network voice terminal meant to support multiple concurrent both two-party and multi-party conferences on top of a variety of transport protocols and using audio encodings offering from vocoder to multi-channel CD quality. As it is to be used as an experimental tool, it offers extensive configuration, trace and statistics options. The design is kept modular so that additional audio encodings, transport and real-time protocols as well as user interfaces can be added readily. In the first part, the report describes the X-based graphical user interface, the configuration and operation. The second part describes the individual components of NEVOT and compares alternate implementations. An appendix covers the installation of NEVOT.

1 Introduction

Increased bandwidth and computational resources have made interactive voice and video communication between workstations across packet communication facilities feasible. Cooperative work, teleconferencing [1] and simple one-to-one "videotelephones" [2, 3] are applications that have attracted a large amount of implementation and research interest.

Transmitting voice and video across a packet-switched network offers a number of advantages other the circuit-switched approach. First, we obtain all the well-known benefits of service integration, particularly important in a multi-media setting. Secondly, we may be able to achieve a higher bandwidth utilization since voice and video do not always use their peak bandwidth (due to silence periods and variable rate coding). Finally, because interleaving several associations tends to be easier in a packet-switched network, control (signaling, to use the telephony term) can be more sophisticated[1].

Research in transmitting voice across a packet network dates back to the early ARPAnet days. Cohen [4] refers to cross-continental packet voice experiments in 1974. According to [5], low-bit

*This work is supported in part by the Office of Naval Research under contract N00014-90-J-1293, the Defense Advanced Research Projects Agency under contract NAG2-578 and a National Science Foundation equipment grant, CERDCR 8500332.

[1]Even narrowband ISDN uses the packet-switched D channel for signaling.

rate voice conferences very carried out in 1976. The early '80s saw experiments of transmitting low-bitrate voice across mobile radio [6, 7] and satellite [8] packet channels. The first Internet packet voice protocol was specified formally in 1977 [9], and a packet video standard followed in 1981 [10]. The CCITT standard G.PVNP [11] was published in 1989. Packet audio/video should be set apart from the approach to voice/data integration that provides fixed-bandwidth circuits on multiple access networks [12, 13].

Interest in packet audio has increased recently as more and more workstations now come equipped with built-in toll-quality (Sun SPARCstations, DEC workstations) or CD-quality (NeXt) audio hardware support. There exist a fair number of simple programs that utilize the SPARCstation audio hardware to communicate between two workstation on a local net, for example vtalk (Miron Cuperman, OKI) or PhoneTalk (Patrik Nises and Joakim Wettby, Royal Institute of Technology, Stockholm). Programs designed for multiple-party connections across wide-area networks include VT [1] and vat (Van Jacobsen and Steve McCanne, LBL). A number of commercial products use medium-bitrate packet voice to more effectively utilize leased private lines, extending the concept of the traditional data-only multiplexer [14]. System implementations of packet voice terminals are described in [5, 18, 19]. Packet radio experiments are featured in [20]. Surveys on packet voice performance are presented in [18].

Numerous other voice/data integration schemes have been studied, usually combining a circuit-switched path for voice and a packet-switched path for data, possibly with bandwidth traded between the two. Examples include [21]. Economic studies comparing alternative network strategies were performed by Gitman and Frank [22].

This report describes NEVOT and is divided into three major parts. The first part, Section 2, describes the facilities of NEVOT and how to use them, principally through the graphical user interface. The second part then delves into the internals, laying out the methods used and comparing some implementation choices. Finally, an appendix provides some hints on installing NEVOT.

2 The Network Voice Terminal (NEVOT) – User's Guide

NEVOT ("NEtwork VOice Terminal") is a tool to support audio conferences across local and wide area networks, including the Internet. It supports multiple simultaneous conferences and a variety of standard and experimental network protocols, including ST-II [23], IP multicast [24, 25, 26] [27, p. 281f] and TCP. It is meant to serve several purposes:

- as a demonstration tool for Internet audio conferences,
- as a measurement tool to investigate traffic patterns and losses in packet voice applications across wide-area networks,
- as a demonstration implementation of real-time services in a distinctly non-real-time operating system (Unix)
- as a traffic source to validate and evaluate resource allocation protocols and algorithms
- as a platform for implementing conference control mechanisms

Extensive tracing and parameterization facilities as well as a modular architecture support experiments in packet voice. The major features are summarized below.

2.1 Features of Version 0.95

Features anticipated for versions released shortly are also listed, but so indicated. Due to operating system or hardware support, a few features are platform-specific. A symbol is used to mark the corresponding platform.

- platforms:
  - Sun SPARCstation¶
  - Silicon Graphics 4D/30 and 4D/35 (Indigo)§
  - Personal DECstation†[in preparation]
- audio protocols:
  - NVP-II (network voice protocol) as used by vat (Lawrence Berkeley Laboratory) and vt (ISI)
  - vat audio packet format
- transport protocols:
  - unicast UDP
  - multicast UDP
  - TCP
  - ST-II¶
- operation as gateway or end system
- compatible with vat session protocol
- user interfaces:
  - XView (OpenLook)
  - Motif GUI
  - curses (for terminals with cursor positioning)
  - dumb terminal
- control:
  - initialization file
  - command line arguments
  - interactive
- several independent concurrent conferences, each with different encoding and compression
- DES-based voice encryption
- current audio encodings supported:
  - 16 bit linear encoding, with all hardware-supported sample rates§
  - 64 kb/s G.711 μ-law PCM

- 32 kb/s G.721 ADPCM¶
- 32 kb/s Intel/DVI ADPCM
- 24 kb/s G.723 ADPCM¶
- 4.8 kb/s LPC (linear-predictive coding) with setable vocoder interval

- dynamic change in audio encoding, with each site having different encodings (but the same sample rate)
- one or multiple audio channels (i.e., mono or stereo)
- playback and recording of audio files (.au and AIFF/AIFC formats), with encoding translation
- extensive statistics and tracing facilities
- arbitrary voice packet length, which may differ for each site
- lost packet substitution
- setable audio buffer occupancy
- configurable adjustment mechanisms for playout delay, VU meter, silence detector and automatic gain control
- redefinable session identifier string with variable substitution

Most commonly, NEVOT interacts with the user through the Open Look™ or Motif™ graphical user interface on X11-capable workstations. Another version with identical functionality, but a more limited user interface, requires only cursor-addressable ASCII terminals supported by the curses library. A fourth version is meant mostly for remote use and uses only terse sequential terminal output to stdio. The command interface used to control the text versions is also available for the XView and Motif versions.

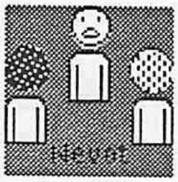

Figure 1: The NEVOT icon

NEVOT may utilize the network services of unicast UDP, multicast UDP, TCP and ST-II [28, 23]. The *source address option* allows operation behind a UDP-level packet reflector, e.g., the simple version written by the author. The packet reflector is used to allow a site running kernels without multicast support to participate in multicast audio conferences. The packet reflector is executed on a multicast-capable site, declares itself part of the multicast group and simply forwards every packet arriving from the multicast group on a unicast UDP socket. The address where the voice packet originated from (i.e., the source address) is prepended by the packet reflector as the first four user data bytes of the packet. This is necessary for proper operation of the voice terminal, since the IP source address of the UDP packet reaching the final destination contains the IP address of the packet reflector rather than that of the speaker, while the actual source address is needed to distinguish several audio streams coming from the same reflector.

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | **sales@docketalarm.com** | **1-866-77-FASTCASE**

fastcase®
Smarter legal research.