Tittel

Forprosjekt om telefoni over Internett

Forfatter

Anders Andersen

<table>
<tr><td colspan="8">Oppdragsgiver
UNINETT</td><td colspan="4">Oppdragsgivers Ref
Prosjeknr 353100.19</td></tr>
<tr><td colspan="3">ISNN 0802-2011
ISBN</td><td colspan="3">Publikasjonsnr
IT.479/1-96</td><td colspan="3">Publikasjonstype
Rapport</td><td colspan="3">Tilgjengelighet
Åpen</td></tr>
<tr><td colspan="4">Dato
5. desember 1996</td><td colspan="4">Versjonsnr
1.1</td><td colspan="4">Antall sider
21</td></tr>
<tr><td colspan="12">Emneord
Internett, kommunikasjon, protokoller, telefoni, multimedia, standarder</td></tr>
<tr><td colspan="12">Noter</td></tr>
<tr><td colspan="12">Distribusjon
Oppdragsgiver</td></tr>
<tr><td colspan="6">Rapportsensor
Asgeir Finnseth</td><td colspan="6">Faglig ansvarlig
Guðmudur Jökulsson</td></tr>
</table>

Sammendrag

UNINETT ønsker å utrede alternativene for å etablere en tjeneste for telefoni i Uninett/Internett. Denne rapporten diskuterer hvordan en slik tjeneste best lar seg realisere med eksisterende teknologi og foreslår to pilotprosjekter med utgangspunkt i denne teknologien.

Forprosjekt om telefoni over Internett

Anders Andersen*
NORUT IT
Anders.Andersen@ITek.NORUT.No

5. desember 1996, versjon 1.1

Sammendrag

UNINETT ønsker å utrede alternativene for å etablere en tjeneste for telefoni i Uninett/Internett. Denne rapporten diskuterer hvordan en slik tjeneste best lar seg realisere med eksisterende teknologi og foreslår to pilotprosjekter med utgangspunkt i denne teknologien.

Innhold

*Dette arbeidet er finansiert av, og gjort på oppdrag for, UNINETT.

1 Innledning

Interessen for telefoni over Internett er meget stor, og ulike leverandører konkurrerer hardt om å få oppmerksomhet omkring sine produkter og standarder. I løpet av det siste året har antall tilgjengelig telefoniprodukter for Internett økt betraktelig. Hovedinntrykket er likevel at produktene ennå er noe umodne og har mange svakheter:

- Lydkvaliteten varierer sterkt.
- Forsinkelser og opphold i lyden er vanlig.
- Kravene til nettkapasitet og endepunkter (datamaskiner) er store.
- Brukervennligheten er meget variabel.
- Ulike produkter *snakker ikke sammen.*
- Standarder er ikke modne og proprietære løsninger er meget utbredt.

Det positive er at vi ser en utvikling hvor disse svakhetene blir gjort noe med. De nyeste produktene har vært et langt steg i riktig retning, og de store aktørene ser ut til å samle seg om et sett av standarder som i det minste delvis kan løse problemet med produkter som *ikke snakker sammen.*

I denne rapporten vil vi i delkapittel 2 se på de ulike formatene som brukes for å digitalisere og komprimere lyd. I delkapittel 3 vil vi se på hvordan den digitaliserte (og komprimerte) lyden kan formidles over nettet, og i delkapittel 4 vil vi se på hvordan forbindelser settes opp og forvaltes. I delkapittel 5 vil vi se på plattform- og systemavhengighet til de ulike løsningene og i delkapittel 6 vil vi diskutere autentisering og avlyttingsprobelmatikk. Etter oppsummeringen følger skisser til pilotprosjekt. Til slutt er det lagt til et tillegg med en oversikt over en del av de produktene som er tilgjengelig i dag og en referanseliste som også inneholder referanser til tilgjengelige ressurser på Internett.

1.1 Notasjon

Vi vil benytte følgende forkortelser i dokumentet:

Kbps	Kilo-biter per sekund, 1024b/s
KHz	Kilo-Herz, 1000Hz
ms	Millisekund, s/1000
PCM	Pulse Code Modulation
DPCM	Differential PCM
ADPCM	Adaptive Differential PCM
SB-ADPCM	Sub-Band ADPCM
SBC	Sub-Band Coding
ATC	Adaptive Transform Coding
AbS	Analysis by Synthesis
MPE	Multi-Pulse Excited
RPE	Regular-Pulse Excited
RPE-LTP	RPE Long-Term Predictor
GSM	Global System for Mobile telecommunication
CELP	Code Excited Linear Prediction
LD-CELP	Low Delay CELP
ACELP	Algebraic Code Excited Linear Prediction
LPC	Linear Predictive Coding
MPEG	Moving Picture Expert Group
CVSD	Continuously Variable Slope Delta Modulation

2 Digitalisering og komprimering av lyd

Vi vil her forsøke å gi en oversikt over og diskutere en del av formatene som brukes for å digitalisere og komprimere lyd. Vi vil konsentrere oss om formater med en viss utbredelse og de riktige egenskapene for Internett telefoni. Formater som er standarder, som har stor utbredelse og som gir et godt forhold mellom kvalitet og kostnad vil bli prioritert.

Vær klar over at bitraten angitt for de ulike kodingene ikke gjenspeiler den toatale mengden av data som overføres i nettet, fordi i tillegg kommer den informasjonen som protokollene for overføring av lyden legger til (hode eller header) [79]. For lyd, som kodes i små pakker, kan størrelsen på denne informasjonen ble betydelig sammenlignet med størrelsen på de dataene som overføres.

For effektiv komprimering og overføring av data bruker mange av metodene å pakke lyden i rammer. Hver av disse kan for eksempel inneholde 20 ms med lyd. Dette introduserer en ekstra forsinkelse i overførelsen av lyden, fordi vi ikke kan spille av noe før en hel ramme er mottatt.

2.1 Metoder for å representere tale digitalt

Før vi går inn på de konkrete formater og standarder vil vi gi en kort innføring i de metodene vi har for å representere og komprimere lyd digitalt [80]. Vi gjør dette fordi det da vil være enklere å forstå og sammenligne de ulike formatene senere.

2.1.1 Bølgekoding

Bølgekoding er et forsøk på kode et signal (lyd) uten å bruke kunnskap om hvordan signalet ble generert. Målet er å kunne rekonstruere signalet så nært opp til det opprinnelige signalet som mulig. Kodere og dekodere for slik koding av lyd er lite komplekse og gir normalt meget god lydkvalitet for ved bitrate over 16 Kbps.

Den enkleste formen for bølgekoding er *PCM* (Pulse Code Modulation), hvor vi med en gitt frekvens måler verdien i aktuelt punkt på bølgeformen som representerer lyden. Med en målefrekvens på 8 KHz dekker målingene frekvensområdet opp til 4 KHz (jmf. Nyquists måleteorem). Ved lineær kvantifisering vil god talekvalitet trenge omkring 12 biter per måling, og dette gir oss et bånbreddekrav på 96 Kbps. Ved bruk av en tilnærming til logaritmisk kvantifisering kan et signal representert ved 8 biter per måling gi et rekonstruert signal som er vanskelig å skille fra originalen. Dette gir oss et båndbreddekrav på 64 Kbps, og brukes i dag både i USA (u-law) og Europa (A-law). Fordelene med PCM er lav kompleksitet, liten forsinkelse og høy kvalitet på den reproduserte talen. Ulempene er relativ høy bitrate og liten toleranse for feil.

Et utbredt teknikk i koding av tale er å forsøke og forutsi neste målte verdi utifra den siste målte verdien. Dette er mulig som en følge av korrelasjonen mellom målingene i tale. En følge av dette er at variasjonen mellom den forutsette verdien og den målte verdien er mindre enn mellom de originale verdiene. Vi kan derfor bruke færre biter enn i det originale signalet for å kvantifisere denne forskjellen. Dette er grunnlaget for *DPCM* (Differential Pulse Code Modulation), hvor vi måler forskjellen mellom det målte og det forventede signalet.

Hvis forventning og kvantifisering lages adaptiv slik at de endres for å tilpasse seg karakteristika til den talen som kodes, så kan vi gjøre ytterlige forbedringer i kodingen av tale. Dette har gitt oss *ADPCM* (Adaptive Differential Pulse Code Modulation). På midten av åttitallet ble ADPCM koding for 32 Kbps standardisert, og senere fulgte også standarder for 16, 24 og 40 Kbps.

Alle metodene for koding som vi har beskrevet ovenfor tar kun utgangspunkt i tidsdomenet. Vi kan også ta utgangspunkt i frekvensdomenet, ved å splitte signalet i flere frekvensband som kodes uavhengig av hverandre med for eksempel en ADPCM koding. Vi kan på den måten allokere flere biter til viktige frekvensbånd slik at disse blir mindre ømfintlig for støy. Dette kaller vi *SBC* (Sub-Band Coding). En variant med adaptive egenskaper er *ATC* (Adaptiv Transform Coding).

2.1.2 Kildebasert koding

Kildebasert koding opererer med en modell av hvordan signalet ble generert. For tale må en slik modell inneholde kunnskap om taleorganene. Ved koding forsøker man å plukke ut parametre til modellen for det signalet som kodes. Disse modellparametrene representerer signalet og brukes for å gjenskape det av en dekoder. Slik koding gir et bitrate på omkring 2.4 Kbps eller mindre. Resultatet er forståelig tale, men som ikke nødvendigvis høres naturlig ut. Det er liten grunn til å forsøke å øke bitraten for å øke kvaliteten ved slik koding, da den innebygde forenklede modellen av tale gir begrensningene på mulig kvalitet.

2.1.3 Hybrid koding

Hybrid koding er et forsøk på å fylle opp mellorommet mellom bølgekoding og kildebasert koding. Kodingen er en hybrid hvor begge de tidligere nevnte metodene benyttes. Det eksisterer ulike typer hybrid koding, men tidsdomene *AbS* (Analysis by Synthesis) kodinger er de mest utbredte og vellykkede metodene. En AbS koding virker ved å splitte opp talen i sett av målinger som kodes i rammer. Hver ramme inneholder 20 ms med tale.

To AbS metoder er *MPE* (Multi-Pulse Excited) og *RPE* (Regular-Pulse Excited). Disse gir god talekvalitet ved bitrate på 10 Kbps og høyere. RPE er noe mer kompleks enn MPE, men gir litt bedre talekvalitet. Det europeiske mobiltelefonsystemet *GSM* benytter en forenklet RPE koding.

Under 10 Kbps er *CELP* (Code Excited Linear Prediction) den mest utbredte kodingen av lyd med god talekvalitet. Den originale CELP kodingen var for kompleks til å bli kodet og dekodet i sann tid. Etter 1985 har det vært arbeidet mye med å redusere kompleksiteten, og i dag er det relativt enkelt å lage en CELP koder og dekoder for sann tid. CELP koding har vært meget vellykket og utbredt for koding av telefonkvalitet tale med en bitrate mellom 4.8 og 16 Kbps.

2.2 Standarder

2.2.1 ITU-TS

ITU-TS[1] (International Telecommunication Union — Telecommunication Sector) har anbefalt en serie med standarder for å komprimere lyd.

G.711 G.711 er simpelthen vanlig PCM. Vi benytter typisk en målefrekvens på 8 KHz og ikke-lineær kvantifisering som krever 8 biter per måling. Dette gir oss en bitrate på 64 Kbps. Talekvalitet vi får er vanskelig å skille fra originalen. I USA brukes u-law koding, og i Europa brukes den nesten like A-law kodingen. Begge disse PCM kodingene er gitt i standarden G.711.

En slik koding er, på grunn av sin enkelhet, gode kvalitet og lave forsinkelse, fortsatt meget utbredt. Såkalte 'au' filer som vi finner i Unix omgivelser eller på Internett er PCM filer.

G.721, G.723, G.726 og G.727 G.721 er en standard for ADPCM koding med en bitrate på 32 Kbps. Denne kodingen gir nesten samme gode kvalitet som 64 Kbps PCM koding. I de senere standardene G.726 og G.727 er også 40, 32, 24 og 16 Kbps ADPCM kodinger spesifisert. G.726 erstatter G.721, og G727 er en utvidelse av denne igjen. G.723, eller mer presis, CCITT G.723, var et subsett av G.726 standarden, og må ikke forveksles med ITU G.723.1 standarden (se nedenfor).

G.722 Målet med G.722 standarden er å få en bedre lydkvalitet enn G.711 og G.721 standardene. G.722 bruker en SB-ADPCM (Sub-Band ADPCM) metode. Denne har en standard målerate på 16 KHz. Mulige bitrater er 48 Kbps, 56 Kbps og 64 Kbps.

G.723.1 G.723.1 (også kalt TrueSpeech 6.3/5.3) [15, 32] har en bitrate på enten 6.3 Kbps eller 5.4 Kbps. G.723.1 benytter en ACELP (Algebraic Code Excited Linear Prediction) metode. I noen tilfeller omtales denne standarden kun som G.723. Den må da ikke forveksles med den gamle CCITT G.723 standarden.

G.728 På bitrater under 16 Kbps faller lydkvaliteten på bølgekodet lyd raskt. Under slike bitrater ser det derfor normalt å bruke hybrid koding som for eksempel CELP. G.728 er en variant av CELP koding, kalt LD-CELP (Low Delay CELP). Vi benytter her kun 5 målinger i en ramme, og dette gir oss en forsinkelse på mindre enn 2 ms. G.728 har en bitrate på 16 Kbps og en forsinkelse på mindre enn 2 ms. Talekvaliteten er den samme som, eller bedre enn, G.721.

2.2.2 GSM 06.10

ETSI (European Telecommunications Standards Institute) spesifikasjon for GSM (Global System for Mobile telecommunication) inneholder en komprimeringsalgoritme for tale som heter *GSM 06.10 RPE-LTP* (Regular-Pulse Excitation Long-Term Predictor) [14], eller kun GSM 06.10. Denne komprimeringen benytter RPE

[1] ITU er tidligere CCITT.

og genererer rammer av 160 13 biters lineære PCM verdier målt ved 8 KHz. Hver ramme dekker 20 ms, og blir komprimert til 260 biter. Dette gir et båndbreddekrav på 13 Kbps. GSM 06.10 gir god talekvalitet, men ikke fullt den samme kvaliteten som G.728.

2.2.3 US Federal Standard 1016

1016 (US Federal Standard 1016) bruker en forenklet versjon av CELP, kalt *LPC* (Linear Predictive Coding). Denne genererer rammer med 30 ms med tale, og har en bitrate på 4.8 Kbps. Resultatet skal være forståelig tale.

2.2.4 MPEG Audio

MPEG (Moving Picture Expert Group) standarden for digitalt video inneholder også en familie med standarder for komprimering av lyd kalt MPEG Audio (noen plasser forkortet MPA).

MPEG-1 Audio MPEG-1 Audio er en familie som består av 3 forskjellige kodinger. Lag 1 er den enkleste, en enkel SBC koder basert på en psykoakustisk modell. I lag 2 har vi en mer avansert bitallokeringsteknikk og en bedre nøyaktighet. Lag 3 legger flere avanserte teknikker som en hybrid filter bank, ikke-uniform kvantifisering, Huffman koding og høyere frekevensoppløsing. Kompleksiteten til algoritmene øker med nummeret på de ulike lagene. De ulike lagene er også hierarkisk kompatibel. Det vil si at en implementasjon av det mest komplekse laget, lag 3, skal kunne kode og dekode lag 1 og lag 2.

MPEG-1 Audio gir mulighet for opptil to lydkanaler (mono, 2 mono, stereo og felles stereo), og målefrekvenser på 48 KHz, 44.1 KHz og 32 KHz. Bitraten kan variere mellom 32 Kbps to 384 Kbps, hvor 96 KHz ansees som meget bra i de fleste tilfeller, og 128 Kbps er for krevende gjengivelser (som pianokonserter).

Lag 2 er identisk med MUSICAM standarden. Den er optimalisert for en bitrate på 96 eller 128 Kbps per monokanal. I stereo er lydkvaliteten nesten identisk med CD-lyd. Lag 1 er en forenklet versjon av MUSICAM med muligheter for enkle kodere og dekodere. Disse vil fungere bra ved 192 eller 256 Kbps. Lag 3 er en kombinasjon av MUSICAM og ASPEC. Målet her er en bitrate på 64 Kbps per lydkanal med en lydkvalitet tett opp til CD, og helt klart bedre enn Lag 2 ved 64 Kbps.

MPEG-2 Audio MPEG-2 Audio støtter 5 full båndbredde kanaler med lyd (surround) og en lavfrekvent tilleggskanal, eller opp til 7 *fortolkende* kanaler. MPEG-2 Audio utvider også stereo og mono kodingen i MPEG-1 Audio med halve målefrekvenser (16 KHz, 22.04 KHz og 24 KHz) for bedret kvalitet for bitrater på 64 Kbps eller mindre per kanal.

2.2.5 Andre

En del andre format for koding av tale er brukt i noen produkter. Vi går ikke inn i detalj på disse fordi de enten er proprietære eller fordi de ikke har stor nok oppslutning blant aktørene. En kort beskrivelse av noen av disse følger likevel nedenfor.

DVI DVI inneholder en ADPCM koding av lyd fra Intel. Vi finner den i flere varianter, hvor DVI bruker 20 ms rammer og har en bitrate på 46 Kbps, DVI2 bruker 40 ms rammer og har en bitrate på 39 Kbps, og DVI4 bruker 80 ms rammer og har en bitrate på 36 Kbps.

LPC4 LPC4 er en LPC koding med 80 ms rammer og en bitrate på 9 Kbps.

CVSD CVSD (Continuously Variable Slope Delta modulation) er en algoritme veldig lik ADPCM. CVSD koding kan operere i området 9.6 til 64 Kbps, men mange implementasjoner er optimalisert for 32 eller 24 Kbps. CVSD er ikke standardisert, og produkter basert på denne koding er derfor proprietær i natur.

WAV WAV (uttales som det engelske ordet *wave*) er utviklet av Microsoft for Windows og DOS maskiner. WAV filer er bølgekodede lydfiler for stereo eller mono, med 8 eller 16 biter per måling. Et stort antall målefrekvenser kan brukes i dette formatet.

AIFF AIFF (Audio Interchange File Format) er utviklet av Apple, men benyttes i dag på flere plattformer. Har mye de samme egenskapene som WAV koding.

RealAudio Progressive Networks har et eget format for koding av lyd [52] som blant annet brukes i deres RealAudio Player. Dette formatet skal kunne spilles av i sanntid over en 14.4 Kbps forbindelse (modem) med bra kvalitet.

2.2.6 Oversikt

Tabellen nedenfor viser en oversikt over standardene for digitalisering og komprimering av lyd som vi har valgt å se på. Kbps angir bitraten. Kvalitet er angitt med utgangspunkt i tre gitte kvaliteter; forståelig tale (tale), standard telefon (toll), og CD-kvalitet (cd). Varianter av disse angis med $+$, $-$ eller $\pm$. Kvalitetsangivelsen har vi kommet frem til fra omtalen av de ulike standardene [78]. De ulike kodingene er altså ikke testet av oss med hensyn på sammenligning av lydkvalitet. Vær også oppmerksom på at kildebasert og hybrid koding vil, på grunn modellene de er bygget opp omkring, kun gi god kvalitet ved tale, og at denne kvaliteten kan variere avhengig av stemmen eller talemåten[2] til personene som snakker.

[2] Det påstås at kvinnestemmer gir dårligere kvalitet enn mannstemmer på det europeiske GSM mobiltelefonnettet.

Navn	Type	Kbps	Kvalitet	Kommentar
G.711	PCM	64	toll	
G.726	ADPCM	16, 24, 32, 40	toll±	Også G.721 og G.727
G.722	SB-ADPCM	48, 56, 64	toll+	
G.723.1	ACELP	6.3, 5.4	toll−	
G.728	LD-CELP	16	toll−	2 ms rammer.
GSM	RPE	13	toll	
1016	LPC	4.8	tale	
MPEG	LPC	32–448	toll – cd+	

3 Formidling av tale

Tale må formidles over nettet i en transportprotokoll. Kravet til en slik transportprotokoll er rimelig store fordi lyd i sin natur er et kontinuerlig medium, og vi hører derfor lett opphold og forsinkelser. Det er ikke store avvik som skal til før lyden er hørbart påvirket. Fordi vi ser på telefoni over Internett, så fokuserer vi på transportprotokoller basert på IP. Dette skaper problemer for oss fordi IP er en datagramprotokoll og foreløpig mangler muligheter for ressursreservering og tjenestekvalitetsgarantier. Vi vil konsentrere oss om en løsning som normalt implementeres over UDP/IP. Dette er en tynn (liten) protokoll kalt *RTP* (Real-Time Protocol). Fordelen med denne protokollen er at den gjør lite. Det vil si at den enkelt kan implementeres over ulike andre transportprotokoller og den kan virke sammen med andre protokoller. Dette kan være protokoller som tar seg av ressursreservering, tjenestekvalitet [5] eller synkronisering [55].

3.1 RTP

RTP (Real-Time Transport Protocol) er en Internett standard for en transport protokoll for sanntids applikasjoner [43, 58, 59, 60]. RTP tilbyr en ende-til-ende transporttjeneste som er egnet for applikasjoner som overfører sanntids data som lyd eller video over mulicast- eller unicastnett. Tjenesten inkluderer typeidentifisering, sekvensnummer og tidsstempel på dataene, og monitorering av overføringen av dataene. RTP adresserer ikke ressursallokering og gir ingen garantier på tjenestekvalitet. Støtte for synkronisering av mediastrømmer er også begrenset.

RTP inkluderer en kontroll protokoll RTCP (Real-Time Control Protocol) som tilbyr en minimal kontroll-, identifikasjon- og overvåkningsfunksjonalitet. RTP og RTCP skal være uavhengige av det underliggende transport- og nettlaget, men tjenestene som tilbys kan påvirkes av egenskapene til disse. RTP støtter dataoverføring til flere mottakere ved multicast hvis dette støttes av det underliggende nettverket.

RTP inneholder også oversettere (translators) og miksere (mixers). Disse gir oss blant annet muligheter til konvertere datastrømmer og blande sammen flere datastrømmer til en RTP-strøm. Slike teknikker kan vi blant annet benytte i forbindelse med brannmurer (firewalls) og nett med ulik båndbredde.

De tidlige versjonene av konferanseapplikasjonen vat (LBL's Audio Conference Tool) brukte en protokoll som senere er blitt referert til som RTP versjon 0. RTP versjon 1 ble publisert i desember 1992, og via flere Internett utkast (Internet drafts) ble RTP versjon 2 i november 1995 godkjent av IESG (Internet Engineering Steering Group) som en foreslått standard.

3.2 Andre

RTSP (Real Time Streaming Protocol) [55] er en applikasjonsnivå-protokoll for å kontrollere overføringen av data med sanntids egenskaper. Protokollen skal kunne kontrollere flere samtidige dataoverføringssesjoner. Formidlingen er basert på RTP. RTSP har fått stor oppslutning blant en del store leverandører [7].

RSVP (Resource ReSerVation Protocol) [5] tilbyr mottakerinitiert ressursreservering for unicast eller multicast dataflyter. RSVP brukes ved forespørsel etter en gitt tjenestekvalitet (QoS) fra nettet til gitte datastrømmer. Protokollen brukes i endepunkter og rutere for å kunne levere og vedlikeholde gitt tjenestekvalitet langs hele ruten til en gitt strøm. RSVP virker over IP (IPv4 og IPv6). RSVP kan benyttes sammen med RTP og RTSP for å kunne tilby bedre garantier for tjenestekvalitet for datastrømmer.

4 Adressering og oppringing

Adressering og oppringing er prosessen med å sette opp en forbindelse mellom to eller flere parter. En forutsetning for å få løst problemet med interoperabilitet mellom produkter er en standardisering av protokollene som er involvert både i adresserings- og oppringningsprosessen og i datautvekslingen. Vi skal først se på noen av problemstillingene omkring dette, og så se på hvorden dette adresseres av noen av standardene.

4.1 Signalering

Ved opprettelse av en forbindelse mellom to (eller flere) parter, må disse bli enige om hvilke type tjenester de skal benytte, kvaliteten på denne, og hvilke format som skal benyttes. En minimumskrav for at to parter skal kunne snakke sammen er en lik forståelse av hvordan signaleringen foregår. Hvis dette kravet er oppfylt, kan forhandlingen om tjenester, kvalitet og formater begynne.

4.2 To- og flerparts samtaler

Toparts samtaler er det spesielle enkle tilfellet. Vi behøver da kun å overføre talen i to strømmer, en i hver retning. Ved flerparts samtaler er problemstillingen betydelig mer kompleks. Ved n deltakere i en telefonkonferanse har vi n talekilder, og n mottakere. Alle mottakere skal kunne motta fra alle kildene (uten seg selv). Vi har her en situasjon med $n-1$ multicast strømmer. For å forenkle dette bildet er det mulig å multiplekse strømmene i en kanal. Støtte for multicast i transportprotokollen vil forenkle implementasjonen av flerparts samtaler. Protokollen RTP støtter multicast hvis de underliggende protokollene gjør dette.

4.3 Katalogtjeneste

Før en forbindelse kan opprettes mellom flere parter må disse ha en metode for å lokalisere hverandre. Skal vi ringe et firma med vanlig telefon, slår vi opp i en telefonliste eller katalog for å finne nummeret. Ved Internett telefoni må vi gjøre noe tilsvarende. Dette kaller vi en katalogtjeneste. En slik katalogtjeneste kan inneholde informasjon om personer, telefonkonferanser eller tjenester, og informasjon om hvordan vi kan kontakte disse med Internett telefon.

4.4 Bro til telefonsystemet

En mulig tjeneste for Internett telefoni er muligheten å ringe abonementer i det vanlige telefonsystemet. Du kan bruke en Internett telefon fra din datamaskin og ringe opp en abonement i det vanlige telefonsystemet ved å kontakte en tilbyder i den aktuelle byen over Internett. Ved en slik samtale vil kostnaden kun være lokaltakst i tilleg til eventuelle kostnader fra aktuell tilbyder. Flere slike tjenester er allerede i drift for noen av Internett telefon produktene.

4.5 Utvidede telefontjenester

En av de viktigste argumentene for å benytte Internett telefoni er de mulighetene dette gir for utvidede tjenester. Dette kan være deling av dokumenter eller tavle (whiteboard) over internett, katalogtjeneste med automatisk signalering, overføring av andre typer data enn tale, og så videre. De fleste Internett telefon produkter tilbyr slike tjenester, men varierer en del på hvilke slike tjenester og hvordan de er integrert i applikasjonen.

4.6 Standarder

Det ser ut til at de fleste store leverandørene har samlet som om et sett av standarder for problemstillingene ovenfor. Den viktigste ser ut til å være H.323 samlingen av standarder som blant annet inneholder signalering, og LDAP for protokolltjenester.

4.6.1 H.323

H.323 [34] beskriver terminaler, utstyr og tjenester for multimediakommunikasjon over LAN (Local Area Networks) som ikke tilbyr tjenestekvalitetsgarantier. H.323 serien inkluderer H.255.0 pakke og synkronisering, H.245 kontroll [33], H.261 og H.263 videokoding, G.711, G.722, G.728, G.729 og G.723.1 lyd koding, og T.120 serien med multimedia kommunikasjonsprotokoller.

Kontrollprotokollen H.245 spesifiserer syntaks og semantikk for informasjonsmeldinger og prosedyrer i forbindelse med oppstart og overføring av data. Det er denne signaleringsprotokollen som gjør det mulig for applikasjoner fra ulike leverandører som snakker H.323 å *snakke sammen.*

4.6.2 LDAP

LDAP (Lightweight Directory Access Protocol) er en katalogtjeneste basert på X.500. LDAP er enklere, lett å utvide og kan bakes inn i ulike typer klienter, tjenere eller applikasjoner. LDAP er basert på kjent Internett teknologi som DNS (Domain Name System) og URL (Uniform Resource Locators).

5 Plattform- og systemavhengihet

Ulike leverandører støtter en eller flere plattformer med sine løsninger. Tendensen til at flere og flere produkter støtter (eller annonserer at de støtter) standarder og protokoller vil kunne fjerne noe av den plattform- og systemavhengigheten vi finner nå.

Microsoft [41, 37, 75] har annonsert at de vil følge standardene RTP, H.323 (inkluderer H.245 og T.120) i NetMeeting applikasjonen. Microsoft ønsker også å benytte RSVP i fremtiden.

Netscape [76, 46] har annonsert at de også vil følge H.323 standarden, og de jobber også mye for RTSP protokollen. LDAP er også en viktig standard for Netscape sine produkter.

Intel [8, 75] følger H.323 standarden i sitt produkt. Intel vil også benytte seg av RSVP og RTP.

Også andre leverandører, som for eksempel VocalTec [19] og RADVision [54] følger opp med støtte for H.323 standarden.

6 Autentisering- og avlyttingsproblematikk

De fleste produkter i dag tar ikke så mye hensyn til autentisering. Kryptering, for å hindre avlytting, er derimot oftere nevnt i produktomtalene. RTP kan tilby støtte for kryptering, men i de fleste tilfeller baserer man seg på kryptering på et lavere lag i nettprotokollen. I RTP kan man også benytte oversettere (translators) for kryptering. En del løsninger har valgt å benytte PGP som krypterings metode.

7 Oppsumering

En del standarder har helt klart pekt seg ut som de mest benyttede i eksisterende og annonserte Internett telefon produkter. For koding av tale ser det ut til at flere alternativer vil bli benyttet. Muligheter for å endre type koding, avhengig av tilgjengelig båndbredde, beregningskapasitet i endepunktene og krav til lydkvalitet og forsinkelse, vil sansynligvis være tilgjengelig i fremtidige produkter. G.723.1 og GSM 06.10 er noen av de formatene som oftest nevnes i produktomtaler, men også andre kodinger har utbredelse.

Overføring av tale over nettet kan løses på mange måter, men på Internett er RTP protokollen den som får mest oppmerksomhet. En bruk av denne over UDP/IP og muligens sammen med andre protokoller kan bli en løsning som vi finner i de fleste produktene. RTSP og RSVP er også protokoller som ofte blir nevnt i forbindelse med økte krav til garantier med hensyn på tjenestekvalitet.

For adressering og oppringing ser det ut til at H.323 serien med standarder tar mer og mer over for de proprietære løsningene vi finner i dag. Vi kan ikke finne andre med tilsvarende oppslutning. Det betyr at vi sannsynligvis vil finne kontrollprotokollen H.245 og samlingen med T.120 protokoller i mange produkter fremover. Om dette betyr at de ulike leverandørenes produkter vil virke godt sammen gjenstår å se.

8 Pilotprosjekt

Vi vil foreslå to mulige pilotprosjekt som en følge av dette forprosjektet. Det første forslaget går ut på å få erfaring med bruk av slike produkter og løsninger ved hjelp av pilotbrukere. Det andre er et prosjekt for å utvikle en plattformuavhengig løsning som støtter de standardene som er foreslått i denne rapporten. Løsningen er tenkt utviklet i Java.

8.1 Utprøvelse av teknologien

Det vil være nødvendig med erfaring med bruk av Internett telefoni produkter i den daglige virksomheten hos pilotbrukere hvis man ønsker en fullverdig evaluering av disse. Dette blir spesielt spennende i tiden fremover hvor mange store leverandører har lovt produkter som skal følge standarder for koding, formidling, signalering og katalogtjenester. I hvor stor grad vil dette føre til at produkter fra ulike leverandører kan virke sammen, og i hvor stor grad blir standardene fulgt opp?

Erfaringene fra slike pilotbrukere kan være en god hjelp for å kunne anbefale produkter og konkretisere de krav man må stille til slike produkter. Hvilke tilleggsfunksjoner på slike produkter er nyttige, og hvilke ønsker til utvidelser har brukerene?

8.2 Plattformuavhengig implementasjon

Ulempen med mange av de produktene som tilbys er at de er tilgjengelig for få plattformer. For en god del av Unix variantene, Mac OS og andre ikke Windows plattformer er utvalget av produkter begrenset. Hvis man ønsker at disse produktene i tillegg skal kunne virke sammen med andre produkter (være basert på åpne standarder), så vil man i mange tilfeller ikke finne noe produkt i det hele tatt.

Vi ønsker å få utviklet en plattformuavhengig implementasjon av en Internett telefon som følger de standardene som industrien ser ut til å velge. For å få til dette ser vi for oss en løsning basert på Java [63, 62].

I første omgang bør man forsøke å få til en begrenset implementasjon som følger de viktigste standardene og som til en viss grad kan virke sammen med andre applikasjoner.

A Produktoversikt

Nedenfor følger en alfabetisk oversikt over en del av de tilgjengelige Internett telefoni produktene.

A.1 Free Phone

Free Phone [17] finnes for ulike varianter av Unix, og er gratis tilgjengelig. Windows 95 versjon vil sansynligvis bli tilgjengelig i løpet av vinteren. Free Phone støtter flere ulike kodingsformater, inkludert GSM 06.10. Free Phone implementerer RTP versjon 2, og skal kunne *snakke sammen* med RAT og vat.

A.2 IBM Internet Connection Phone

IBM Internet Connection Phone [22] for Windows variantene bruker stort sett egen teknologi. Tale kodes med en modifisert GSM metode.

A.3 Intel Internet Phone

Intel Internet Phone [24] får vi for Windows 95. Intel Internet Phone følger H.323 standarden, og skal derfor i teorien kunne *snakke sammen* med andre produkter som følger denne standarden.

A.4 Microsoft NetMeeting

Microsoft NetMeeting [40] følger med versjon 3 av Internet Explorer. Microsoft går for en løsning basert på H.323, G.723.1, RTP og RSVP. RTSP blir også vurdert.

A.5 Netscape CoolTalk

Netscape (InSoft) CoolTalk [45] følger med Netscape Navigator 3.0. CoolTalk er tilgjengelig for Windows variantene, Mac OS og en rekke Unix varianter.

A.6 NeVoT

NeVoT [48] er tilgjengelig for ulike Unix varianter (Sun, SGI). NeVoT bruker vat og RTP protokollene over UDP, og støtter flere ulike kodinger av tale.

A.7 Speak Freely

Speek Freely er tilgjengelig for Unix [73] og Windows variantene [74]. Speek Freely støtter RTP protokollen, og skal kunne kommunisere med vat kompatible program.

A.8 VocalTec Internet Phone

VocalTec Internet Phone [68] er tilgjengelig for Windows 3.1, Windows 95 og Mac OS.

A.9 Andre

Andre produkter som er tilgjengelig er ClearPhone [6], DigiPhone [64], FreeTel [18], IRIS Phone [31], Net2Phone [23] (bro til telefonsystemet), OnLive! Talker og Community Server [50], PGPfone [51], RAT [56], SoftFone [61], TeleVox [72], vat (LBNL Audio Conferencing Tool) [36], WebPhone [47] og WebTalk [53].

B Referanser

[1] **Andersen, Anders**. *Forprosjekt om telefoni over Internett.*
`http://www.itek.norut.no/~anders/dist/doc/uninett/rapport/rapport.html`
`http://www.itek.norut.no/~anders/dist/doc/uninett/rapport.ps.gz`
Prosjektrapport (denne rapporten), NORUT IT, November 1996.

[2] **Atlanta Signal Processors, Inc**. *An Introduction to Speech Coding.*
`http://www.aspi.com/products/coder.htm`
1991, Revised July 94.

[3] **Baldazo, Rex and Mathog, Michael**. *Web Phones.*
`http://www.cnet.com/Content/Reviews/Compare/Webphone/index.html`
Comparative Reviews, CNET 1996.

[4] **Bormann, Carsten**. *Providing integrated services over low-bitrate links.*
`http://www2.es.net/pub/internet-drafts/draft-ietf-issll-isslow-00.txt`
Internet-Draft June 1996.

[5] **Braden, R. and Zhang, L. and Berson, S. and Herzog, S. and Jamin, S.**. *Resource ReSerVation Protocol (RSVP) — Version 1 Functional Specification.*
`ftp://ietf.org/internet-drafts/draft-ietf-rsvp-spec-14.txt`
`ftp://ietf.org/internet-drafts/draft-ietf-rsvp-spec-14.ps`
Internet-Draft. November 5, 1996.

[6] **ClearPhone**. *Home.*
`http://www.kaiwan.com/~radiobob/`

[7] **Communicationsweek**. *Streaming Branches on the 'Net — Protocol vs. Microsoft's proprietary technology in push to ease multimedia access.*
`http://www.inquiry.com/publication/cmp/CWK/19961028/CWK19961028S0121.html`
October 28, 1996.

[8] **Davis, Jim**. *New line of Net phones .*
`http://www.news.com/News/Item/0,4,3703,00.html`
CNET News.Com, September 20.

[9] **December Communications, Inc**. *Summary of Internet Tools.*
`http://www.december.com/net/tools/about.html`
This summary of Internet tools is a collection of information sources about software used on the Internet for network information retrieval, computer-mediated interaction, and computer-mediated communication.

[10] **December Communications, Inc**. *CMC.*
`http://www.december.com/net/tools/cmc.html`
Computer-Mediated Communication.

[11] **December Communications, Inc**. *Standards.*
`http://www.december.com/net/tools/standards.html`
Protocols and standards are the basis for operating tools and forums on the Internet.

[12] **Decker, Jack**. *Audio and Video via the Internet.*
`http://www.novagate.com/~jack/audiovid.html`
Links to applications and resources.

[13] **Defense Information Systems Agency (DISA)**. *Multimedia Technology Standards Assessment.*
`http://OTT.sc.ist.ucf.edu/refs/stds/mtsa/`
Version 2, August 1995, Center for Standards.

[14] **Degener, Jutta**. *Putting the GSM 06.10 RPE-LTP algorithm to work.*
`http://www.ddj.com/ddj/1994/1994.12/degener.htm`
Dr. Dobb's Journal, December 1994.

[15] **DSP Group, Inc**. *ITU G.723 Speech Compression Technology for the ITU H.324 and ITU H.323 Videoconferencing/Telephony Standards.*
`http://www.dspg.com/tsg723.htm`

[16] **ESCAtech media inc**. *Computer Sound Formats.*
`http://www.escatech.com/soundpage/sound.html`
About computer sound formats, including examples.

[17] **FreePhone**. *Free Phone — Telephony over the Internet.*
`http://www.inria.fr/rodeo/fphone/`

[18] **FreeTel Communications**. *FreeTel — Talk over Internet for free!.*
`http://www.freetel.inter.net/`

[19] **Galante, Suzanne**. *Short Take: VocalTec embeds open standard into products.*
`http://www.news.com/News/Item/0,4,4965,00.html`
CNET News.Com, October 30, 1996.

[20] **Hertz Technologies Inc.**. *The Home Of Hertz Technologies Inc.*
`http://www.hertz1.com/`
Presents some of the Internet software products available anywhere on the net.

[21] **Hertz Technologies Inc.**. *Real time voice communication over the internet.*
`http://www.hertz1.com/voice.htm`
List of products for real-time voice communication over the Internet.

[22] **IBM**. *Internet Connection Phone.*
http://www.internet.ibm.com/icphone/index.html

[23] **IDT**. *Net2Phone.*
http://www.net2phone.com/

[24] **Intel**. *Intel Internet Phone.*
http://www.intel.com/iaweb/cpc/iphone/index.htm

[25] **International Teleconferencing Association**. *Home.*
http://www.itca.org/
International Teleconferencing Association Homepage.

[26] **International Teleconferencing Association**. *Multimedia Standards Update.*
http://www.itca.org/pubs/connections/9604.html
April 1996.

[27] **Internet Telephony Consortium**. *Home.*
http://itel.mit.edu/
A research organization focused on providing interoperability between the Internet and the public switched telephone network.

[28] **Internet Telephony Consortium**. *TC Directory Services Study Group.*
http://itel.mit.edu/directory.html
The goal of the Directory Services Study Group of ITC is to help resolve interoperability problems with directory services for Internet telephony.

[29] **Internet Telephony Consortium**. *Speech Compression/DSP Resources.*
http://itel.mit.edu/dsp.html
Links to resources about speech compression.

[30] **Internet Telephony Consortium**. *Internet Telephony Standards.*
http://itel.mit.edu/standards.html
The Internet Telephony Consortium list of standards.

[31] **IRIS Systems, Ltd**. *IRIS Phone.*
http://irisphone.com/

[32] **ITU-T**. *Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s.*
http://www.itu.ch/itudoc/itu-t/rec/g/g700-799/g723-1_32261.html
ftp://ftp.std.com/vendors/PictureTel/h324/g723_rev.doc
DRAFT ITU-T Recommendation G.723.1.

[33] **ITU-T**. *Control protocol for multimedia communication.*
ftp://standard.pictel.com/avc-site/H245NCM6.DOC
ftp://standard.pictel.com/avc-site/H245NCM6.DOC.zip
DRAFT ITU-T Recommendation H.245.

[34] **ITU-T**. *Visual Telephone Systems and Equipment for Local Area Networks which Provide a Non-Guaranteed Quality of Service.*
ftp://standard.pictel.com/avc-site/H323NCM3.DOC
ftp://standard.pictel.com/avc-site/H323NCM3.DOC.zip
DRAFT ITU-T Recommendation H.323. May 28, 1996.

[35] **Jade**. *Internet Phones.*
http://www.is.unimt.mt/~jade/iphone.html
Software, hardware and other resources.

[36] **Lawrence Berkeley National Laboratory**. *Vat - LBNL Audio Conferencing Tool.*
http://www-nrg.ee.lbl.gov/vat/

[37] **MacDonald, Christine**. *Microsoft takes a meeting.*
http://www.news.com/News/Item/0,4,4414,00.html
CNET News.Com, October 14.

[38] **Marz**. *Internet Voice Products Review.*
http://ourworld.compuserve.com/homepages/presdee/voice.htm
Test of some of the Internet phone products.

[39] **Microsoft**. *Home.*
http://www.microsoft.com/
Microsoft Homepage.

[40] **Microsoft**. *Microsoft NetMeeting.*
http://microsoft.com/netmeeting/
Microsoft NetMeeting Homepage.

[41] **Microsoft**. *Microsoft NetMeeting Conferencing Software Provides Easy Voice, Data Internet Communications.*
http://www.microsoft.com/corpinfo/press/1996/may96/INCONFPR.HTM
Microsoft Press Release about NetMeeting.

[42] **NACSE**. *Analysis of current MBone Re-play Techniques.*
http://www.CS.ORST.EDU/~kleinro/nacse/replay/
Contains a report on a study of currently existing tools and techniques for re-playing of recorded visual and audio MBone material.

[43] **NACSE**. *The Realtime Transport Protocol.*
http://www.CS.ORST.EDU/~kleinro/nacse/rtp/
Contains a report on a study of the Realtime Transport Protocol (RTP) in combination with the Realtime Transport Control Protocol (RTCP).

[44] **Netscape**. *Homepage.*
http://home.netscape.com/

[45] **Netscape**. *Cooltalk.*
http://home.netscape.com/comprod/products/navigator/version_3.0/communication/cooltalk/index.

[46] **Netscape**. *Netscape Conference.*
http://home.netscape.com/comprod/products/communicator/guide.html#conference

[47] **NetSpeak Corporation**. *Homepage.*
http://www.itelco.com/

[48] **NeVoT**. *Guide to NeVoT.*
http://www.cs.columbia.edu/~hgs/nevot/

[49] **NORUT IT**. *Hjemmeside NORUT Informasjonsteknologi.*
http://www.itek.norut.no/itek/
Prosjektet er utført av NORUT IT, November 1996.

[50] **OnLive!**. *Homepage.*
http://www.entertv.com/

[51] **Price, Will**. *PGPfone - Pretty Good Privacy Phone.*
http://web.mit.edu/network/pgpfone/

[52] **Progressive Networks**. *RealAudio.*
http://www.realaudio.com/

[53] **Quarterdeck**. *WebTalk.*
http://arachnid.qdeck.com/qdeck/products/webtalk/

[54] **RADVision**. *RADVision Announces H.323 Software Stack.*
http://www.radvision.com/pr.htm#1
August 12, 1996.

[55] **Rao, Anup and Lanphier, Rob**. *Real Time Streaming Protocol (RTSP).*
ftp://ds.internic.net/internet-drafts/draft-rao-rtsp-00.txt
Internet-Draft.

[56] **RAT**. *The RAT (Robust-Audio Tool) Home Page.*
http://www-mice.cs.ucl.ac.uk/mice/rat/

[57] **Savetz, Kevin M. and Sears, Andrew**. *How can I use the Internet as a telephone?.*
http://rpcp.mit.edu/~asears/voice-faq.html
http://www.northcoast.com/~savetz/voice-faq.html
FAQ, 1995–1996.

[58] **Schulzrinne, Henning**. *Real-Time Transport Protocol (RTP).*
http://www.cs.columbia.edu/~hgs/rtp/
Department of Computer Science, Columbia University, 1996.

[59] **Schulzrinne, H. and Casner, S. and Frederick, R. and Jacobson V.**. *RTP: A Transport Protocol for Real-Time Applications.*
ftp://ftp.ds.internic.net/rfc/rfc1889.txt
http://www.cs.columbia.edu/~hgs/rtp/rfc1889.ps.gz
Request for Comments: 1889.

[60] **Schulzrinne, H.**. *RTP Profile for Audio and Video Conferences with Minimal Control.*
ftp://ftp.ds.internic.net/rfc/rfc1890.txt
http://www.cs.columbia.edu/~hgs/rtp/rfc1890.ps.gz
Request for Comments: 1890.

[61] **SilverSoft**. *SoftFone.*
http://www.pak.net/softfone.htm

[62] **Skrivervik Data and Sun Microsystems**. *Java.*
http://www.java.no/

[63] **Sun Microsystems**. *Java*.
`http://www.sun.com/java/`

[64] **Third Planet Publishing**. *Digiphone*.
`http://www.digiphone.com/digiphon/digindex.htm`

[65] **Uninett**. *UNINETTs web-tjeneste*.
`http://www.uninett.no/`
Oppdragsgiver i dette prosjektet.

[66] **Uninett**. *MICE*.
`http://sauce.mmlab.uninett.no:80/mice-nsc/`
Multimedia Integrated ServiCes for European Researches.

[67] **Uninett**. *Sanntids Multimedia Stttesenter*.
`http://www.hiof.no/smm/`
Uninett støttesenter for PC-basert sanntids multimediakommunikasjon.

[68] **VocalTec**. *Internet Phone Release 4*.
`http://www.vocaltec.com/iphone4/ip4.htm`

[69] **Voice on the Net**. *VON*.
`http://www.von.com/`
Audio, video and telephony on the net; products and resources by Jeff Pulver.

[70] **Voice on the Net**. *About Voice on the Net*.
`http://www.von.com/about.html`
Comments about audio, video and telephony on the net by Jeff Pulver.

[71] **Voice on the Net**. *Internet Telephony*.
`http://www.von.com/teleph.html`
Internet Telephony products and resources by Jeff Pulver.

[72] **Voxware**. *TeleVox*.
`http://www.voxware.com/televox/moretelevox.html`

[73] **Walker, John**. *Speek Freely for Unix*.
`http://www.fourmilab.ch/speakfree/unix/`

[74] **Walker, John**. *Speek Freely for Windows*.
`http://www.fourmilab.ch/speakfree/windows/`

[75] **Wingfield, Nick**. *MS, Intel make conferencing deal*.
`http://www.news.com/News/Item/0,4,1835,00.html`
CNET News.Com, July 17, 1996.

[76] **Wingfield, Nick**. *Netscape opens intranet attack*.
`http://www.news.com/News/Item/0,4,4462,00.html`
CNET News.Com, October 15, 1996.

[77] **Winkler, Jeff**. *A/V Conferencing Overview*.
`http://www.nlanr.net/COLL/VTC/vconf.html`
Wide Area Network Multimedia Collaboration, 1995.

[78] **Winkler, Jeff**. *Some Bit Rate and Compression Information.*
http://www.nlanr.net/COLL/VTC/rateinfo.html
Wide Area Network Multimedia Collaboration, 1995.

[79] **Winkler, Jeff**. *Audio Rate Comparisons.*
http://www.nlanr.net/COLL/VTC/audio_rates.html
Wide Area Network Multimedia Collaboration, 1995.

[80] **Woodard, Jason**. *Speech Coding.*
http://www-mobile.ecs.soton.ac.uk/speech_codecs/
Department of Electronics & Computer Science, University of Southampton.