Sikkerhet i Voice over IP og andre multimediasesjoner basert på SIP og RTP

Anne Pernille Hveem

Forsvarets forskningsinstitutt (FFI)

27. juni 2012

FFI-rapport 2012/00521

1242

P: ISBN 978-82-464-2105-6
E: ISBN 978-82-464-2106-3

Emneord

VoIP

SIP

SIPS

Sikkerhet

Godkjent av

Kjell Olav Nystuen	Prosjektleder
Anders Eggen	Avdelingssjef

Sammendrag

”Voice over IP” (VoIP), også kjent som IP-telefoni, er telefoni over pakkebaserte nettverk, slik som Internet. Hovedårsakene til utbredelsen av VoIP er at en unngår å ha et eget nett for formidling av tale og at VoIP kan tilby mer funksjonalitet enn tradisjonell ”Public Switched Telephone Network” (PSTN). De siste 10 årene har det vært en nedgang i bruken av tradisjonell linjesvitsjet PSTN, samtidig som det har vært en jevn økning av VoIP. Undersøkelser fra 2011 viser at 31% av fasttelefonmarkedet i Norge benytter VoIP.

De vanligste åpne protokollene for transportering av VoIP og andre multimediasesjoner i dag er ”Session Initiation Protocol” (SIP) og ”Real-time Transport Protocol” (RTP). SIP benyttes for signallering, mens RTP benyttes for selve samtalen (mediasesjonen). Selv om andre protokoller kan tilby VoIP, har SIP og RTP blitt *de facto* industristandard for VoIP. Det finnes også en god del lukkede proprietære VoIP-løsninger som er veldig store, blant annet Skype og Google Talk. Skype og Google Talk har gode sikkerhetsløsninger med god kryptering. Denne rapporten vil omhandle VoIP realisert med SIP og RTP.

Målet med rapporten er å gi en oversikt over de viktigste sårbarhetene i og diskutere sikkerhetsproblematikk for VoIP og andre multimediasesjoner basert på SIP og RTP. I likhet med vanlig telefoni er VoIP sårbart for flere reelle trusler og angrep. Gode sikkerhetsmekanismer er foreløpig ikke tatt i bruk i noen vesentlig grad i SIP og RTP basert VoIP, blant annet på grunn av økt kompleksitet ved bruk, mangelfull støtte i VoIP-produktene og at få vil betale for å få dekket sikkerhetstjenester. Rapporten kommer inn på sikkerhetsmekanismer som kan brukes for å sikre SIP-signallering og RTP-mediastrøm.

Det har tradisjonelt vært lite fokus på sikkerhet i SIP, som viser seg blant annet ved en sårbar autentisering. SIP og RTP basert VoIP har dessuten vist seg sårbar for blant annet ulovlig avlytting og trafikkanalyse, endring av VoIP-trafikkdata (tale og signallering) og DoS-angrep.

Sikring av VoIP er imidlertid mer enn bare å implementere sikkerhetsmekanismer i SIP og RTP. Siden sikkerhetsdesignet i SIP og RTP er såpass fraværende, anbefales det at man benytter seg av andre sikkerhetsmekanismer som implementerer ønskede sikkerhetsegenskaper (konfidensialitet, integritet osv). Dette kan være oppsett av IPsec-tunneler mellom VoIP-leverandører, ”Virtual Private Network” (VPN) mot endebrukere, brannvegger og så videre.

Det finnes flere sikkerhetsmekanismer for SIP-signallering som vil bli diskutert i denne rapporten, blant annet Secure SIP (SIPS) og S/MIME. Det viser seg dessverre at disse, og andre sikkerhetsmekanismer, har fått ingen eller svært begrenset utbredelse. SIPS har imidlertid fått en del oppmerksomhet de siste årene og det jobbes en del med SIPS både i standardiseringsorganet IETF, og hos VoIP-leverandører. Tiden vil vise om SIPS, eller andre sikkerhetsmekanismer, vil vinne frem. Standardprotokollen for beskyttelse av sanntids multimediakommunikasjon som tale og video er ”Secure Real Time Protocol” (SRTP). SRTP-protokollen er ikke mye brukt i dag, men også den har fått en del oppmerksomhet og støtte i industrien i den senere tid.

English summary

Voice over IP (VoIP), also known as IP telephony, is telephony over packet switched networks, like the Internet. Increased functionality and no need to operate a separate network for voice are the main reasons VoIP has gained popularity. There has been a steady increase of VoIP users the last ten years coincident with a decrease in Public Switched Telephone Network (PSTN) users. Research from 2011 shows that around 31% of the fixed telephone market in Norway use VoIP.

The most common open protocols for transporting VoIP and other multimedia sessions today are Session Initiation Protocol (SIP) and Real-time Transport Protocol (RTP). SIP is used for signaling, while RTP is used for the content (media session). Even though there are other protocols which offer VoIP, SIP and RTP have become de facto industry standard. There are also several proprietary solutions for VoIP that take a good portion of the VoIP-market, for instance Skype and Google Talk. Skype and Google Talk have good security solutions with good encryption. This report will look into VoIP based on SIP and RTP.

This report aims at giving an insight into vulnerabilities in and discussing security issues for VoIP and other multimedia sessions based on SIP and RTP. SIP and RTP based VoIP is today vulnerable to a number of critical attacks with dire consequences. High quality security mechanisms have currently seen low industry penetration, due to increased complexity for said security mechanisms, lack of implementation in the VoIP products and lack of funding for deploying secure VoIP installations. This report looks into security mechanisms which can be used to secure SIP signaling and RTP media session.

Traditionally it has been little focus on security design in SIP. The authentication in SIP, for example, is known to be very vulnerable. SIP and RTP based VoIP is also vulnerable to eavesdropping and traffic analysis, change in both media stream and signaling and different kinds of DoS attacks.

Securing VoIP is not only about implementing security mechanisms for SIP and RTP. Since the security design is almost absent in SIP and RTP, use of other security mechanisms to implement the desired security properties (confidentiality, integrity et cetera) is recommended. This can be setup of IPSec-tunnels between VoIP vendors, Virtual Private Network (VPN) towards end users, firewall et cetera.

Other security mechanisms for securing SIP signaling that will be discussed in this report are among others Secure SIP (SIPS) and S/MIME. Unfortunately these and other security mechanisms have gained none or very limited industrial adaption. SIPS has however got some attention lately and the standardization body (IETF) and VoIP vendors are now working with SIPS. Time will show if SIPS or other security mechanisms will gain popularity. The standard protocol for protecting real time multimedia communication like voice and video is Secure Real Time Protocol (SRTP). The SRTP protocol is not commonly used today, but also SRTP has got some industrial support lately.

Innhold

Forord

Jeg vil sende en stor takk til Lars Strand for hans gode innspill og konstruktive tilbakemeldinger i mitt arbeid med denne rapporten.

1 Innledning

Det pågår i dag en omfattende utvikling innen offentlig elektronisk kommunikasjon. I dette bildet er det blant annet en klar trend at militære og sivile infrastrukturer smelter sammen. Sivile teknologier vil dermed i stadig større grad ha betydning for Forsvaret. I den forbindelse vil det som ledd i arbeidet med 1126 UNET på FFI gjøres et arbeid med VoIP realisert med protokollene SIP og RTP. Denne rapporten påpeker en rekke sårbarheter ved SIP- og RTP-basert VoIP i dag og kommer inn på sikkerhetsmekanismer som kan brukes for å sikre SIP- og RTP-basert VoIP i fremtiden.

1.1 Målsetting med rapporten

Det har vært en markant økning av VoIP-brukere samtidig med en nedgang i PSTN-brukere de siste 10 årene. På midten av 2011 var andelen som benyttet VoIP med fasttelefoni i Norge 31% i følge Post- og teletilsynet [1]. Antallet som benytter VoIP er antagelig langt høyere, da tallene ikke tar hensyn til andre bruksformer for VoIP som PC-til-PC (Skype), ”softphones”, online gaming, VoIP-klienter på mobiltelefoner etc.

Realisering av VoIP-telefoni i dag utgjør et broket landskap med mange forskjellige løsninger. De største åpne standardene for realisering av VoIP er SIP og RTP. Det finnes i tillegg mange proprietære løsninger for realisering av VoIP som er blitt veldig store, som for eksempel Skype og Google Talk. Skype ble registrert med 663 millioner brukere på verdensbasis september 2011. De er også store på private utenlandssamtaler, og spesielt populær blant ungdommer [2]. Da det er vanskelig å tilegne seg informasjon om protokollene i proprietære løsninger vil denne rapporten fokusere på de åpne standardene SIP og RTP for realisering av VoIP.

Målet med rapporten er å gi en oversikt over en rekke sårbarheter ved SIP og RTP basert VoIP i dag, samt diskutere sikkerhetsproblematikk i VoIP og andre multimediasesjoner basert på SIP og RTP. Rapporten vil i første omgang se på SIP- og RTP-protokollene, hvor SIP håndterer signalleringen og RTP selve innholdet av sesjonen (tale/video). Andre del av rapporten har fokus på sårbarheter og sikkerhetsmekanismer ved VoIP basert på SIP og RTP.

1.2 Rapportens oppbygning

Kapittel 2 gir en generell innføring i VoIP. VoIP-protokollene SIP og RTP blir belyst i kapittel 3. Kapittel 4 tar for seg sårbarheter og sikkerhetsmekanismer i VoIP og andre multimediasesjoner realisert med SIP og RTP. Oppsummering og konklusjon følger i kapittel 5.

2 Voice over IP (VoIP) og multimediasesjoner

Teknologien for implementering av ”Voice over IP” (VoIP) og andre multimediasesjoner, benytter seg av pakkesvitsjet IP-nett (slik som Internet) for leveranse av tale og andre multimediasesjoner som for eksempel video. Pakkebaserte nettverk kan støtte både data og

sanntid multimediaapplikasjoner som VoIP eller Video over IP. Dette kapittelet gir en generell innføring i VoIP.

2.1 VoIP-egenskaper

VoIP muliggjør at taletjenester kan benytte eksisterende IP-nett istedenfor et eget linjesvitsjet PSTN-nett. Det å slippe å operere flere nett og muligheten til å legge til flere tjenester er i dag hovedårsakene til utbredelsen av VoIP hos tjenesteleverandørene. Mange VoIP-selskaper tilbyr gratis samtaler mellom abonnenter i samme selskap, og for en flat månedlig avgift tilbys ofte ubegrenset ringetid innenlands. Noen ganger inkluderer denne avgiften også ubegrenset ringetid til noen andre utvalgte land.

I tillegg til talekommunikasjon kan VoIP blant annet tilby [3]:

- Video-samtaler
- Forbedret talekvalitet (”HD sound”)
- Fax
- Roaming
- Konferanse mellom flere enn to aktører
- Integrering med andre applikasjoner som for eksempel ”Instant Messaging” (IM), epost og kalender. Når VoIP integreres med e-post vil talebeskjeder konverteres til epost og sendes til mottager og omvendt. VoIP integrert med IM gjør det mulig for bruker å ringe fra IM-brukergrensesnittet. VoIP integrert mot annen infrastruktur som IM, epost og kalender kalles med et samlebegrep ”Unified Communication” (UC).

Mens PSTN-telefoner er fysisk koblet til en linje, så er VoIP-telefoner uavhenging av fysisk plassering. Derfor kan et VoIP-telefonabonnement brukes hvor som helst i verden. Dette gjør det vanskeligere å lokalisere den som ringer ved nødanrop, da IP-adresser ofte er dynamiske og det er vanskelig å stedfeste dem geografisk.

Siden endeutstyret hos VoIP-abonnenten ofte er beskyttet bak en brannmur kan abonnenten ikke bli oppringt uten å skreddersy brannmuren til å holde en port åpen for IP-pakker fra VoIP-leverandøren. Om ikke VoIP-tjenesten leveres med egen ruter som inneholder en brannmur, er dette vanskelig å få til hos alle kunder som ønsker en enkel installasjon. Derfor bruker VoIP-leverandører en løsning hvor abonnentens terminal hele tiden sender en meldig (hver 15 sekund for Telio) og spør etter anrop og på denne måten holder en port i brannmuren åpen.

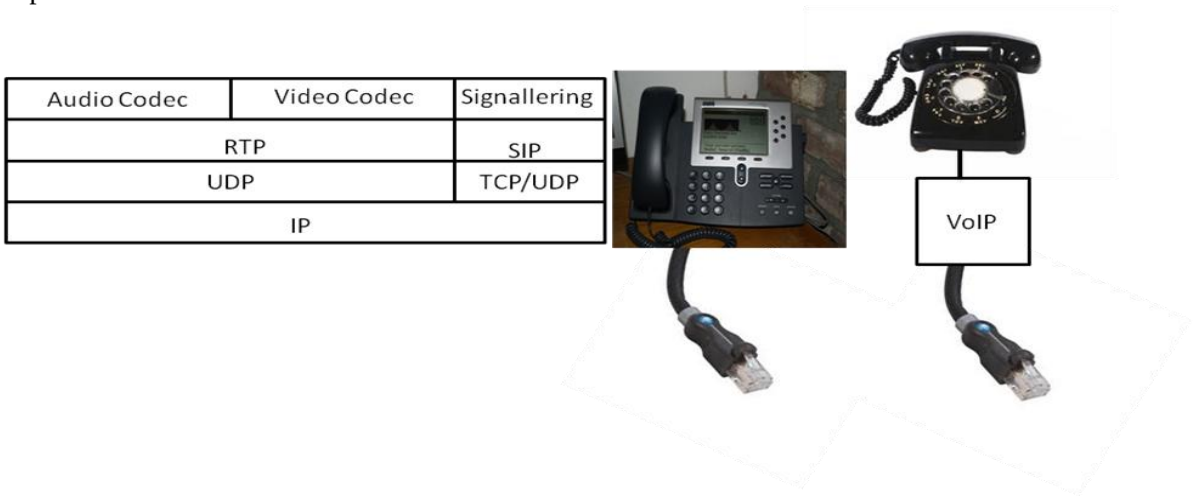
VoIP-telefoni er sårbar for feil på strømforsyningen, da den ikke har egen reserve-strømforsyning slik som de opprinnelige telefonapparatene i PSTN.

2.2 VoIP-arkitektur

Nytt utstyr må benyttes for å tilkoble VoIP. Telefonen må støtte VoIP, alternativt kan en konverter, ”Analog Telephone Adapter” (ATA), kobles til for å oversette ”gammeldags” telefoni

(PSTN) til VoIP. Det er også mer og mer vanlig å benytte seg av programmer på PC som tilbyr VoIP, såkalt ”softphone”, slik at man er fristilt fra å anskaffe dedikert utstyr.

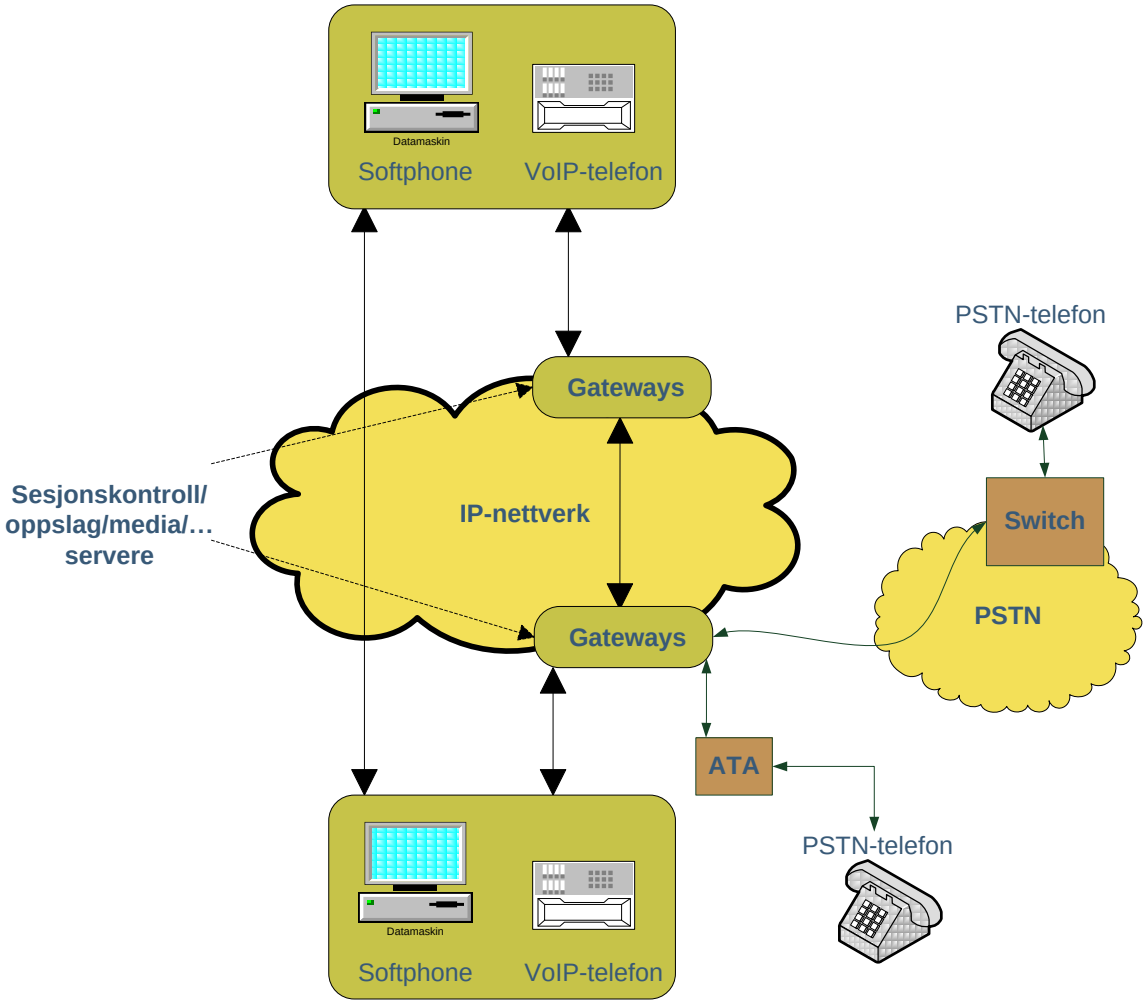
Figur 2.1 viser hvorledes gamle telefoner kan kobles til et IP-nett via en enhet fra en VoIP-leverandør, eller ved en telefon som støtter VoIP. Enheten inneholder en prosessor og en Analog –Digital / Digital-Analog konverter. En tale-, video-codec plasserer henholdsvis tale- og video-pakker inn i IP-pakker ved bruk av ”Real-time Transport Protocol” (RTP) og ”User Datagram Protocol” (UDP). Signalleringen skjer via ”Session Initiation Protocol” (SIP) via ”Transmission Control Protocol” (TCP) eller UDP og IP. Se protokollstakken i Figur 2.1. Enheten kobles til datanettet via en vanlig Ethernet-kabel. En VoIP-telefon kan også koble seg til IP-nettverk ved bruk av trådløst nettverk; IEEE 802.11. For mer informasjon om SIP og RTP-protokollene se kapittel 3.



Figur 2.1 Tilkobling av endeutstyr til IP-nett via VoIP-leverandør

IP-pakkene fra endeutstyret adresseres til VoIP-leverandørens gateway som igjen oftest er koblet mot det offenlige telefonnettet (PSTN) og IP-nettverk (som kan være Internet). Siden signallering og tale er svært forskjellig i et IP-nett og PSTN, utfører VoIP-leverandørens gateway konvertering i begge retninger.

Det eksisterer ikke en standardisert VoIP-arkitektur som kan dekke alle mulige installasjons-scenarier og funksjonalitet i VoIP. Likevel er det mulig å referere til en generell VoIP-arkitektur for å belyse funksjonelle krav og komponenter. Se Figur 2.2 for et eksempel på en generell VoIP-arkitektur. Figuren viser kommunikasjon mellom to VoIP-endepunkter over IP-nettverk (som kan være Internet) eller mellom et VoIP-endepunkt og PSTN. Kommunikasjon over IP-nettverk går enten peer-til-peer eller via VoIP-leverandørens gateway. En gateway vil rute signallering og tale mellom forskjellige nettverk og blant annet oversette signallering og transkode tale der det er nødvendig. Som for eksempel hvis to IP-nettverk har forskjellig signallering og ved overgang fra IP-nettverk til PSTN. En gateway vil også kunne tilby funksjonalitet som autentisering og adgangskontroll av VoIP-samtaler. De to VoIP-endepunktene i Figur 2.2 kan ha samme VoIP-leverandør og da vil samtalen rutes innenfor det lokale IP-nettverket til VoIP-leverandøren.



Figur 2.2 VoIP-arkitektur

2.3 VoIP-ytelse

For at VoIP skal bli et levedyktig alternativ til linjesvitsjet PSTN, må det også kunne tilby like god eller bedre talekvalitet til kunden. Kommunikasjon over IP-nettverk er generelt ikke like pålitelig som over PSTN. Det eksisterer ingen mekanismer i IP-nettverk for å sikre seg at datapakker ikke mistes eller kommer frem i riktig rekkefølge. Det finnes ingen ”Quality-of-Service” (QoS) garantier, men QoS-mekanismer som øker sannsynligheten for levering av pakker innen gitte tidsfrister.

Oppsett av en VoIP- eller multimedia-sesjon involverer følgende steg:

1. Initiere og forhandle frem kontekstinformasjon for en samtale/sesjon (signallering) mellom to eller flere parter. Dette utføres av protokollen SIP.
2. Digitalisering av tale/video, samt pakke dette ned i egnet codec-format (eks:G.721)
3. Sende sesjonsinnholdet mellom de kommuniserende parter ved hjelp av RTP.
4. Terminere sesjonen ved hjelp av SIP

Kodek (Koder-dekoder) er en tale- eller videokompresjonsmekanisme, som komprimerer antall bit som blir brukt til å representere tale- eller videosignalet. Årsaken til at kompresjonsmekanismene blir benevnt kodeks er fordi det komprimerte signalet er forståelig bare for dekoder og kan derfor bli sett på som om det var kodet. Kodeks er definert gjennom ulike standarder som beskriver algoritmer for hvordan taleinformasjon skal komprimeres og pakkes. RTP-protokollen bestemmer videre hvordan taleblokkene fra koder skal plasseres i en eller flere IP-pakker for sending til mottager. En talekoder reduserer mengden av informasjon i signalet som skal overføres og dekoder gjenoppretter signalet på mottagersiden. Det dekodede signalet vil ikke være eksakt lik det opprinnelige signalet og forskjellen utgjør støyen som er blitt tilført signalet av kompresjonsalgoritmen. Kompresjonsalgoritmene utnytter at mennesker tolererer en god del støy hvis den blir tilført på riktig måte. Hvilken kodek som blir brukt varierer veldig. Noen kodeks er optimert på båndbredde, mens andre er optimert på talekvalitet og tilbyr høykvalitets stereolyd.

God talekvalitet er avhengig av at forsinkelsen ikke blir for stor. Talepakkene bæres fram av IP-pakker som benytter RTP. Når forsinkelsen øker, må deltagerne som kommuniserer lære seg å ikke avbryte hverandre og vente før man prater. Om taleforsinkelsen er mindre enn ca. 20ms legger man ikke merke til den [4]. For å forhindre lang forsinkelse benytter RTP-protokollen transportprotokollen UDP, hvor det ikke er noen retransmisjon av pakker. Dette betyr at pakker kan gå tapt. For tale og video oppleves det ikke sjenerende når bare en eller to IP-pakker blir borte. Når dette skjer blir bare de siste mottatte IP-pakkene gjentatt.

Veien IP-pakkene går gjennom Internet bestemmes for hvert enkelt hopp av rutere. IP-pakkene legges i kø i en buffer på valgt ruterutgang, hvor fordeleren velger hvilken pakke som skal sendes ut på linken etter ”First In First Out” (FIFO) prinsippet. Behovet for buffer kommer av at utgangslinken ikke har kapasitet til å sende all trafikk den får fra ruterinngangene. Hvis bufferet er fullt vil innkommende pakker bli droppet. Bufring av IP-pakken tilfører ekstra forsinkelse og jo flere rutere IP-pakken må gjennom jo større forsinkelse får den.

Kølengden i ruteren kan variere over tid avhengig av trafikkbelastning. Ruteren svitsjer IP-pakker fra samme samtale individuelt, slik at de kan gå forskjellige veier gjennom IP-nettverket. Variasjon i kølengde og veivalg betyr at IP-pakker fra samme samtale får ulik forsinkelse, denne variasjonen i forsinkelse kalles jitter. Hvis IP-pakkene spilles av i rekkefølgen de ankommer mottager vil stor jitter føre til at pakkene enten kommer tett eller langt fra hverandre i tid. For å motvirke denne degraderingen av talekvalitet brukes et jitterbuffer hos mottager. Jo større jitterbuffer jo bedre kan nettverket klare å redusere effekten av jitter. For å unngå for stor forsinkelse i VoIP kan jitterbuffer ikke bli veldig stort [3].

VoIP-codecs har mekanismer som kan håndtere pakketap og jitter. ”Adaptiv Multirate codecs” som opererer på flere bitrater og som kan tilpasse seg forsinkelse og tilgjengelig båndbredde er tilgjengelig i dag. Det utvikles også stadig nye VoIP-codecs for å oppnå bedre robusthet til variasjoner i nettverkstilstand og derav bedre talekvalitet i heterogene IP-nettverk.

Da Internet ikke tilbyr noen QoS-garanti har ”the Internet Engineering Task Force” (IETF) foreslått mekanismer hvor trafikk med sanntidskrav kan bli prioritert av ruterne for å motvirke pakketap, stor forsinkelse og jitter. Disse mekanismene heter; ”Integrated Services” (IntServ) og ”Differentiated Services” (DiffServ), hvor DiffServ er mulig å realisere i stor skala. En utfordring ved bruk av DiffServ er at flere ”Internet Service Providers” (ISPs) må samarbeide på trafikk eller pakkenivå for å kunne gi den ønskede QoS-garantien [3]. Dessverre blir verken IntServ eller DiffServ benyttet i noe stor skala i industrien og de fleste anser denne teknologien som uaktuell.

”Multiprotocol label Switching” (MPLS) kan benyttes for å reservere kapasitet i nettet. Ved bruk av MPLS vil datapakker bli tildelt en merkelapp, og bli sendt gjennom ruterne i en dedikert kanal. Det arbeides en del med MPLS i dag.

3 VoIP-protokoller

Den mest brukte åpne signalleringsprotokollen for VoIP og andre multimedisasesjoner i dag er ”Session Initation Protocol” (SIP). SIP har tatt over for den eldre signalleringsprotokollen H.323. Det eksisterer i tillegg som tidligere nevnt en god del proprietære protokoller, blant annet Skype, som blir mye brukt for VoIP. Protokollspesifikasjonen og kildekoden i Skype er lukket og det er derfor vanskelig for utenforstående å se hvordan protokollene i Skype fungerer. Denne rapporten vil kun ta for seg SIP, selv om H.323 fortsatt er støttet og i bruk. Den mest brukte åpne protokollen for transport av tale og video er ”Real-time Transport Protocol” (RTP). Selv om andre protokoller kan tilby VoIP, har SIP og RTP blitt *de facto* industristandard for VoIP. Innholdet i kapittelet er hovedsaklig hentet fra [3;5-9].

3.1 Session Initiation Protocol (SIP)

”Session Initiation Protocol” (SIP) er designet og standardisert av ”Internet Engineering Task Force” (IETF) [9]. SIP er en signalleringsprotokoll på applikasjonslaget som blir brukt til å sette opp, endre og avslutte multimediasesjoner. Sesjonene kan bestå av en eller flere mediastrømmer mellom to eller flere parter.

SIP er bygget opp rundt en forespørsel og respons modell, hvor den ene siden sender en forespørsel og den andre siden responderer på forespørselen. Hver SIP-enhet har mulighet til å gjøre begge deler avhengig av hvilken enhet som initierer utvekslingen. SIP-meldingene er i ren tekst som gjør at de kan leses av utenforstående. SIP-protokollen er uavhengig av det underliggende transportlaget og kan bruke både transportprotokollen ”User Datagram Protocol” (UDP) og ”Transmission Control Protocol” (TCP), men bruker vanligvis UDP. To SIP-endepunkter kan kommunisere uten noen mellomliggende infrastruktur, derfor blir den ofte beskrevet som peer-to-peer. Dette er imidlertidig ikke gunstig for tjenestetilbydere da de blant annet ikke kan ta betalt for tjenesten.

Det er også sjelden at den som skal ringe kjenner IP-adresse eller vertsnavn til mottaker, i tillegg til at disse kan endres ved mobilitet, noe som krever en SIP-proxy for å rute signallering. Hovedkomponentene i en SIP-arkitektur er:

- *User Agent (UA)*
- *SIP-Registrar*
- *SIP-proxyserver*
- *SIP- gateway*

En ”User Agent” (UA) er en logisk komponent i en SIP-sesjon som kan sende eller motta SIP-meldinger. En VoIP-telefon vil inneholde en UA som i en SIP-sesjon vil sende eller motta SIP-meldinger avhengig av om VoIP-telefonen ringer ut eller får en innkommende samtale.

SIP-Registrar er en logisk komponent hvor bruker autentiserer og registrerer sin identitet og lokasjon. Dette skjer når telefonen slås på, skifter IP-adresse eller nettverk og ved jevne tidsintervall (5-60min). SIP-Registrar er vanligvis integrert i en SIP-proxyserver.

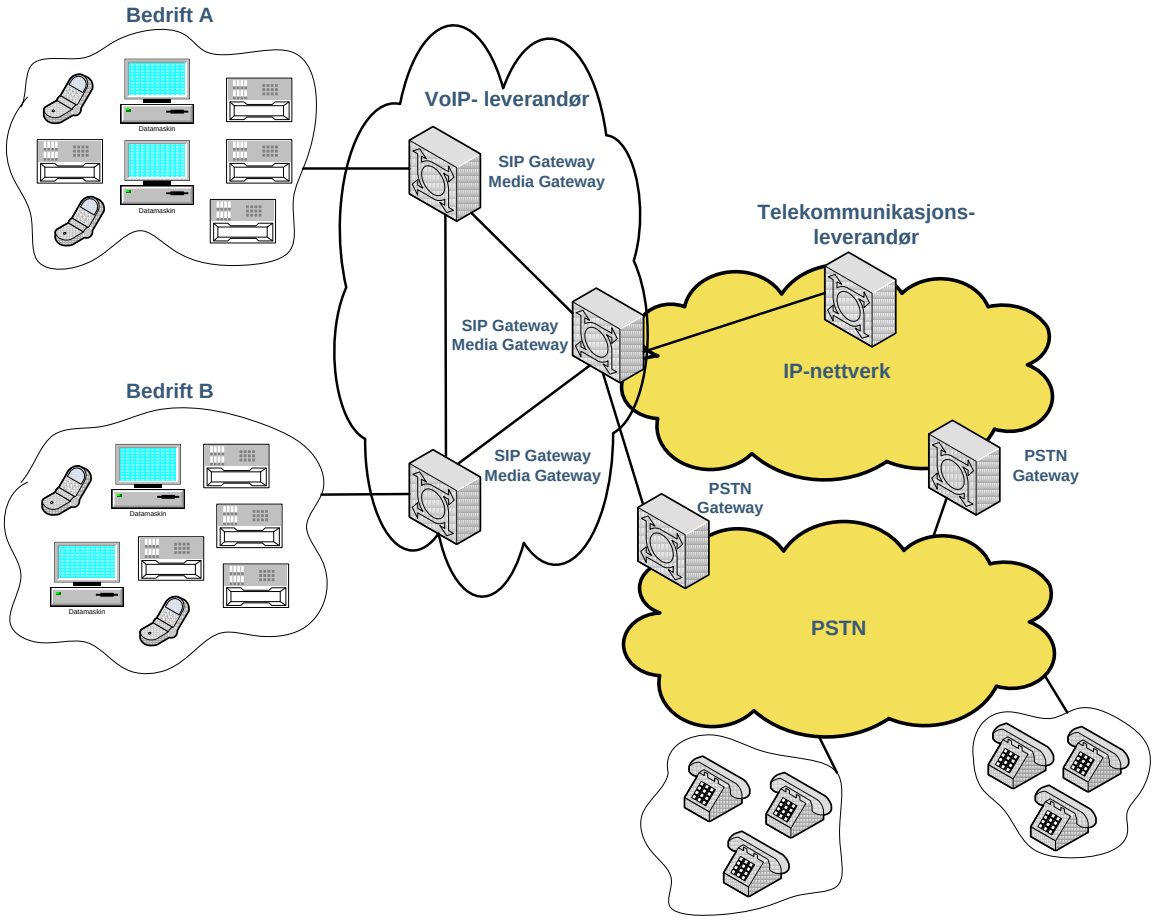
SIP-proxyserver er en logisk komponent som videresender signalleringsmeldinger mellom UAene. Når SIP-proxyserver mottar en SIP-melding vil den rute meldingen i retning mottaker-UA. SIP-proxyserver ruter og videresender samtaleforespørselen til mottaker-UA og returnerer responsen til UAen som initierte samtaleanropet. En SIP-proxyserver kan derfor ses på som en ruter for signalleringstrafikken. SIP-proxyserver kan også operere som en SIP-gateway.

SIP-gateway er en logisk komponent som sjekker og ruter SIP-meldinger ut og inn fra ett domene. Domene er en fysisk eller logisk avgrensning, gjerne ett bestemt nettverk eller en bedrift.

I tillegg har vi to logiske komponenter som håndterer mediastrømmen; RTP-proxy og mediagateway. RTP-proxy er en mediaserver som videresender mediastrømmen mellom UAer. En RTP-proxy kan operere som en mediagateway når den står på grensen av et domene, og rute trafikk inn og ut av domenet.

En SIP-bruker adresseres ved hjelp av en SIP ”Uniform Resource Identifier” (URI). SIP kan bruke både tall og tekst som ”telefonnummer”. En typisk URI kan derfor være sip:22123456@corp.com eller sip:bob@corp.com med brukernavn@domenenavn som ved email. Hensikten med email-modellen er at en bruker skal kunne ringe en annen bruker direkte uten å gå via en eller flere VoIP-leverandører, såkalt global adressering. E.164 NUmber Mapping (ENUM) standarden var et forslag for å tilby global adressering. ENUM skisserer opp en måte for å ”oversette” telefonnummer til ”Domain Name System” (DNS). Man skal da kunne bruke vanlig DNS-oppslag til å slå opp telefonnummer og annen relevant VoIP-informasjon. Adressering ved bruk av email-modellen har ikke tatt av i industrien og det er mest vanlig i dag å bruke vanlige telefonnummer. Noen av grunnene til dette er at VoIP-leverandører ønsker å tjene penger på å rute hverandres trafikk, i tillegg til en rekke sikkerhetsutfordringer ved bruk av email-modellen, som blant annet uønskede samtaler ”Spam over Internet Telephony” (SPIT), autentisering av bruker og DoS-angrep [10].

Her følger et eksempel på hvordan vanlige telefonnummer brukes i VoIP i dag. En lokal bedrift med 100 ansatte har satt opp en VoIP-løsning, hvor de bruker vanlige (PSTN) nummer (8 siffer). Da det er en middels til stor bedrift, klarer de seg med én fysisk server. På denne er det installert for eksempel Asterisk, som er en software-implementasjon av en bedrifts telefonsentral. Asterisk tar seg av både SIP- og RTP-funksjonalitet. All VoIP-trafikk som ikke er lokalt mellom ansatte, rutes ut til en VoIP-leverandør som håndterer alle eksterne telefonsamtaler. Asterisk er satt opp til å bruke vanlige telefonnummer til adressering via SIP. Når for eksempel Kari ringer Ola på telefonnummer 22 12 34 56, så vil Asterisk se at Ola er en lokal bruker og søker opp IP-adressen til Ola i sin database. Asterisk vil så sende SIP-meldingen til denne IP-adressen. SIP URIen kan da se ut som sip:22123456@corp.com eller sip:22123456@localdomain. Hvis telefonsamtalen går ut av bedriften havner den hos VoIP-leverandøren som bedriften har en avtale med (for eksempel Ventelo). Her vil VoIP-leverandøren ha en stor database med "sine" telefonnummer, som aldri rutes ut av leverandørens (Ventelos) nett. En stor VoIP-leverandør kan rute VoIP-trafikk mellom ulike domener uten å gå ut på andre IP-nettverk eller PSTN ved å bruke lokal ruting innad i eget IP-nett. Hvis ikke telefonnummeret finnes hos leverandøren, rutes den til riktig leverandør som har dette telefonnummeret, for eksempel Telenor, Netcom eller en annen VoIP-leverandør. Se Figur 3.1 for et eksempel på et VoIP-scenario hvor to bedrifter er knyttet opp mot en VoIP-leverandør.



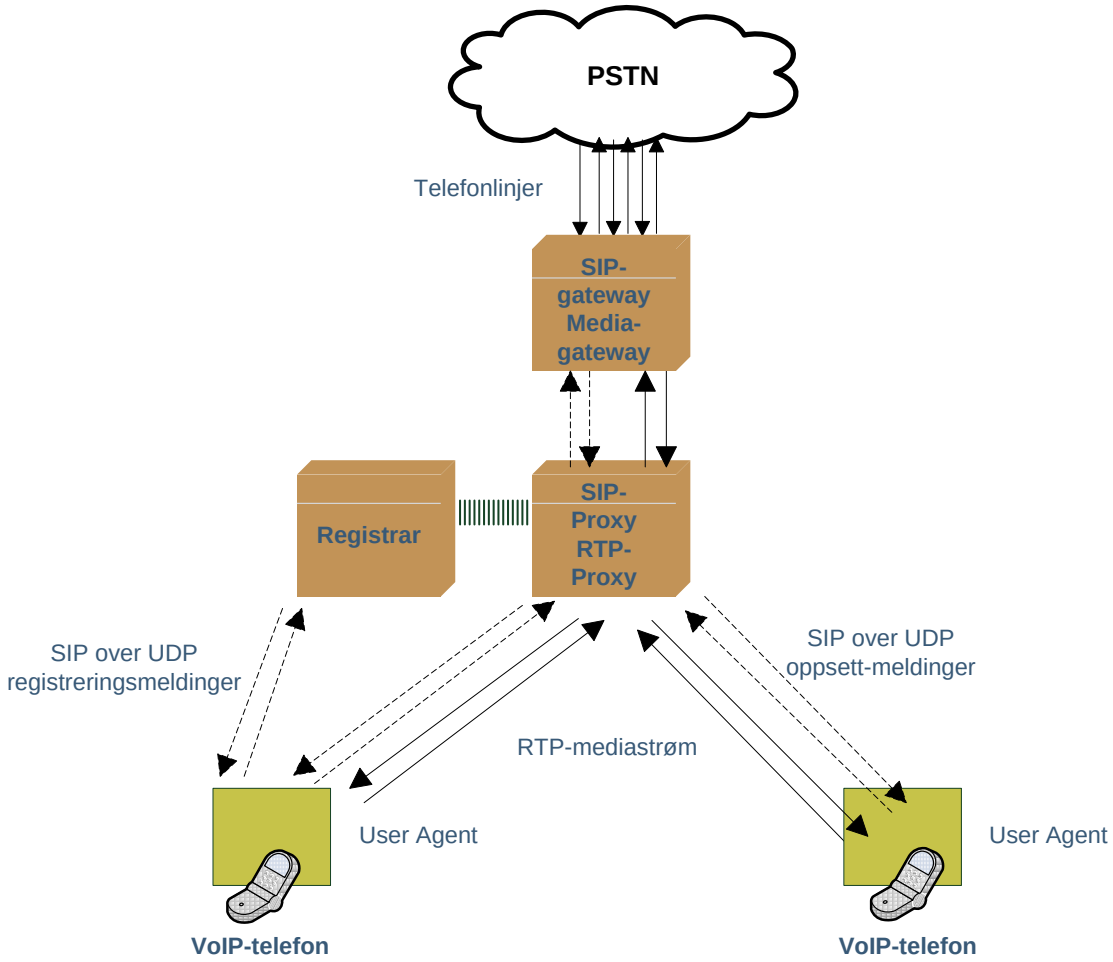
Figur 3.1 Arkitektur for VoIP-scenario

Mediaforbindelsen som SIP-signalleringen setter opp, kan enten gå direkte mellom sender og mottaker (peer-til-peer) eller via RTP-proxy (mediaserver) eller mediagateway. Det vanligste er at RTP-trafikken går gjennom en eller flere mediagateway på veien til mottaker. En fysisk server kan rute både SIP-signalleringstrafikk og RTP-mediastrøm, og vil da bestå av en SIP-proxy og en RTP-proxy. Når den fysiske serveren står på grensen av ett domene vil SIP-proxy og RTP-proxy fungere som henholdsvis en SIP-gateway og mediagateway og rute signallering og trafikk inn og ut av domenet. Det vil også eksistere konfigurasjoner hvor en SIP-gateway og mediagateway kommer i tillegg til SIP-proxy og RTP-proxy. VoIP gir et sett med byggeklosser med ulik funksjonalitet. En systemarkitekt kan bygge sin egen installasjon basert på krav og ønsker, hvor de viktigste momentene er skalerbarhet, størrelse og sikkerhet. VoIP-installasjoner vil derfor variere avhengig av behov. Situasjoner hvor mediaforbindelsen må gå via mediaserver og eventuelt mediagateway:

- når UAene ikke har direkte tilgang til hverandre, for eksempel når en eller begge UAer befinner seg bak en ”Network Adress Translation” (NAT)/brannmur.
- ved registrering av varighet av samtale (ved for eksempel taksering etter varighet)
- ved krav til lovlig sporing av samtale
- ved krav til lovlig avlytting av samtale (”lawful interception”)
- når bruker kommer til en talemeny og må taste tall for å komme videre i menyen, så må disse tallene registreres av mediaserver

På grunn av operatørkrav til lovlig avlytting av samtale vil i praksis all kommersiell kommunikasjon gå via mediaserver og eventuelt mediagateway. Det finnes allikevel VoIP-leverandører som omgår disse kravene.

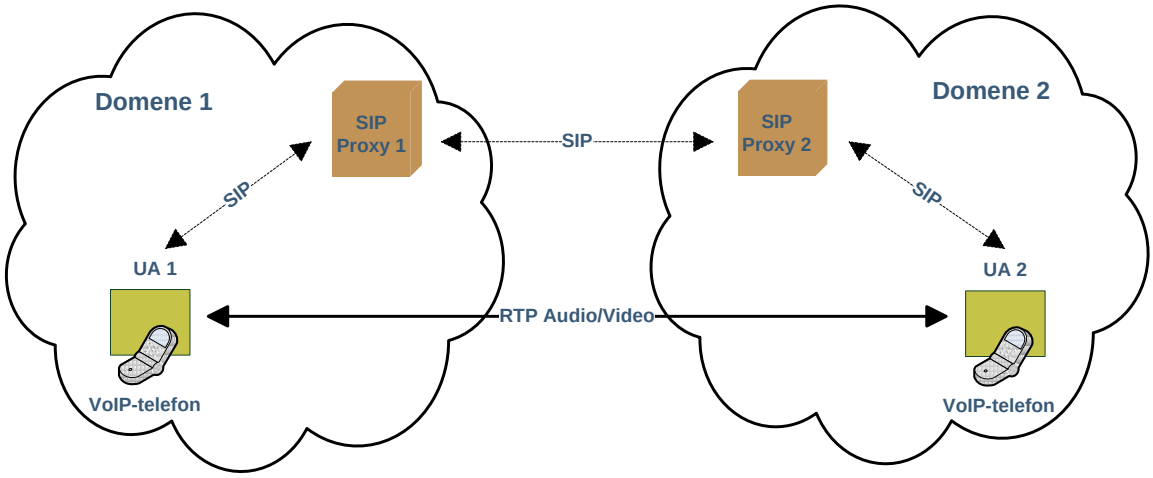
Figur 3.2 viser et eksempel på en SIP-arkitektur hvor RTP-sesjonen går via RTP-proxy. Ved lokale samtaler innad i et domene kan sesjonen settes opp direkte via proxyen. Hvis samtalen går ut av domene til for eksempel PSTN må signallering og mediastrøm rutes ut ved hjelp av SIP-gateway og RTP-gateway (mediagateway). En RTP-sesjon består av en gruppe deltagere som kommuniserer over RTP. De stiplede pilene i figuren viser SIP-signallering og de heltrukne pilene viser RTP-mediastrøm. Figuren viser også SIP-Registrar som tar seg av registrering og som oftest også autentisering av bruker. All lokasjonsdata lagres i SIP-Registrar.



Figur 3.2 VoIP-arkitektur hvor RTP-trafikk går via RTP-proxy [5]

For å forhandle frem konteksten av en VoIP-samtale eller multimediasesjon, benyttes ”Session Description Protocol” (SDP). SDP benyttes for å forhandle frem sesjonsparametre som mediatype og -format, IP-adresse/vertsnavn, portnummer m.m. Selve SDP blir transportert som nyttelast (”payload”) i SIP-meldinger. Blant annet ved SIP-INVITE meldinger benyttes SDP. Mer detaljer om SDP i [11].

Figur 3.3 viser kommunikasjonsutveksling mellom to UAer som tilhører forskjellige domener, hvor mediastrømmen går direkte mellom UAene. SIP-signalleringen går via SIP-proxyene. Det kan også være flere SIP-proxyservere mellom de to domenene som signalleringen går igjennom. Når SIP-proxy 2 får en forespørsel, vil den kontakte lokasjonsdatabasen til domene 2 for å finne IP-adressen assosiert med URI til UA2 og sende forespørselen videre til UA2.



Figur 3.3 SIP-arkitektur med to domener, hvor RTP-trafikk går direkte mellom UAene

De vanligste SIP-meldingsforespørsler er:

- REGISTER – registrering eller avregistrering av identitet og lokasjon til bruker
- INVITE – initierer eller reforhandler en multimediasesjon
- ACK – bekrefter en suksessfull etablering av en sesjon
- BYE – avslutter en multimediasesjon
- CANCEL – kansellering av en forespørsel

De vanligste SIP-responsene er:

- 100 Trying – SIP-proxyserver jobber med å etablere kontakt med ønsket UA
- 180 Ringing – det ringer hos ønsket UA
- 200 OK – forespørselen er mottatt, forstått og akseptert av ønsket UA
- 401 Unauthorized
- 407 Proxy Authentication Required

En SIP UA responderer med en eller flere SIP-responser på en SIP-forespørsel. Flest responser (2xx, 3xx, 4xx, 5xx, 6xx) er endelige og avslutter den gjeldende SIP-meldingsutvekslingen, mens 1xx responsene er midlertidige og avslutter ikke SIP-meldingsutvekslingen. For mer detaljert informasjon om SIP-forespørsler og SIP-responser se [7;9].

Eksempel på en SIP-INVITE melding er vist i Figur 3.4. Denne SIP-meldingen er en del av en utgående samtale initiert av VoIP-bruker "Alice" med destinasjon til "Bob" innenfor samme domene "corp.com". Meldingen kan deles inn i tre:

1) Første linje, markert i blå, viser hvilken SIP-metode benyttet (INVITE) og hvem som er mottager av meldingen.

2) Neste gruppe, markert i lys grønn, er selve SIP-headerne, som inneholder kontekst-informasjon om SIP-meldingsutvekslingen. For eksempel inneholder Via-header hvilken IP-adresse og UDP-port 5060 som skal benyttes for å sende svar på SIP-invitasjonen. Content-Type-header viser at SIP-meldingen har en payload (SDP).

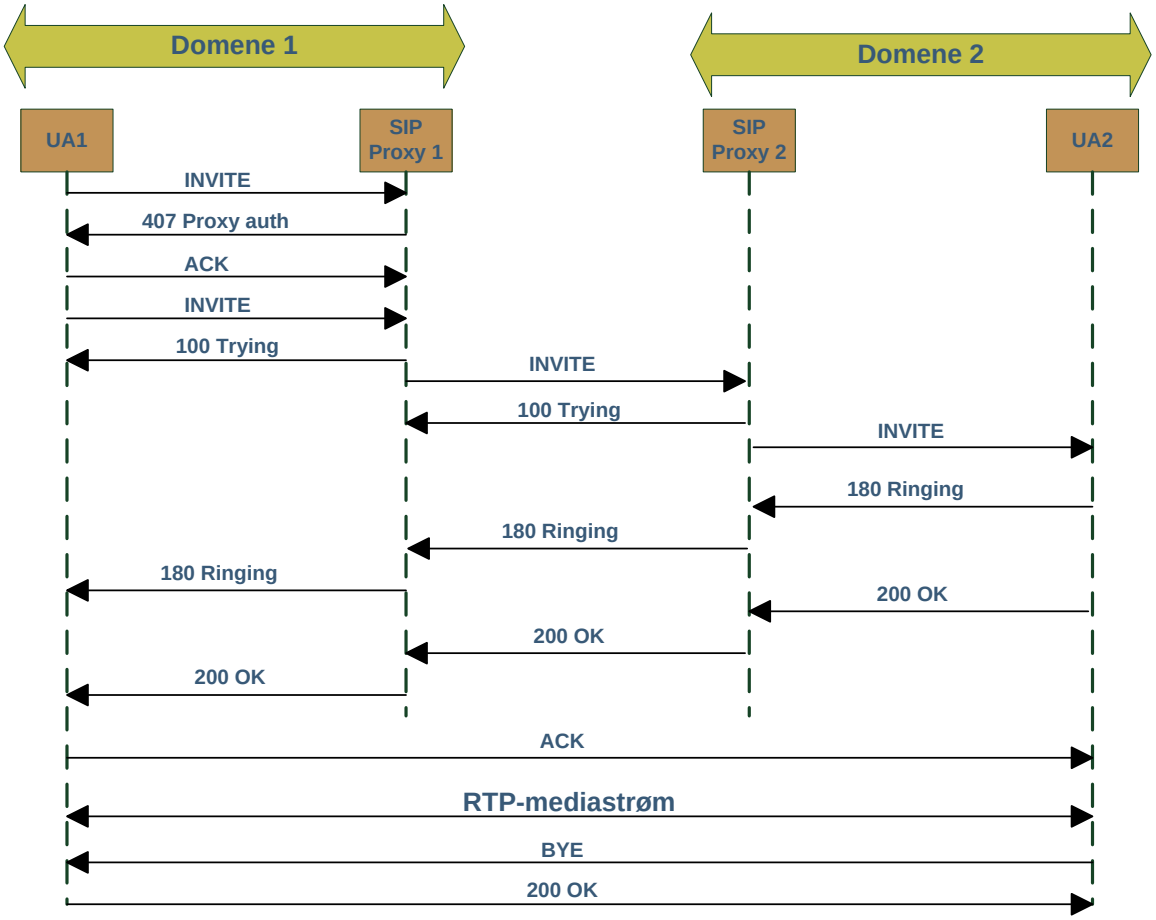
3) Siste gruppe, markert i gult, er SIP payload (SDP). Denne payloaden er en del av informasjonsutvekslingen som er nødvendig for å sette opp en RTP-strøm og inneholder kontekst-informasjon om RTP-strømmen (selve samtaleinnholdet). SDP-nyttelasten i Figur 3.4 gir blant annet informasjon om senders IP-adresse, at det er en audiosamtale hvor telefonen vil lytte på UDP-port 9000 (ukryptert), og hvilke tale codec som telefonen støtter [5].

```
INVITE sip:Bob@corp.com SIP/2.0
Via: SIP/2.0/UDP 192.168.0.10:5060;branch=z9hG4bK922648023
From: ”Alice”<sip:Alice@corp.com>;tag=1687298419
To: <sip:Bob@corp.com>
Supported: replaces, 100rel, timer
Call-ID: 2114455679@192.168.0.10
CSeq: 20 INVITE
Session-Expires: 1800
Contact: <sip:Alice@192.168.0.10>
Max-Forwards: 70
Expires: 180
Content-Type: application/sdp
Content-Length: 217
v=o
o=Alice 1352822030 1434897705 IN IP4 192.168.0.10
s=A conversation
c=IN IP4 192.168.0.10
t=0 0
m=audio 9000 RTP/AVP 0 8 18
a=rtpmap:0 PCMU/8000/1
a=rtpmap:8 PCMA/8000/1
a=rtpmap:18 G729/8000/1
a=ptime:20
```

Figur 3.4 Eksempel på en SIP-INVITE melding [5]

Figur 3.5 viser SIP-meldingsutveksling for oppsett og terminering av en RTP-sesjon direkte mellom to UAer i forskjellige domener (som i Figur 3.3):

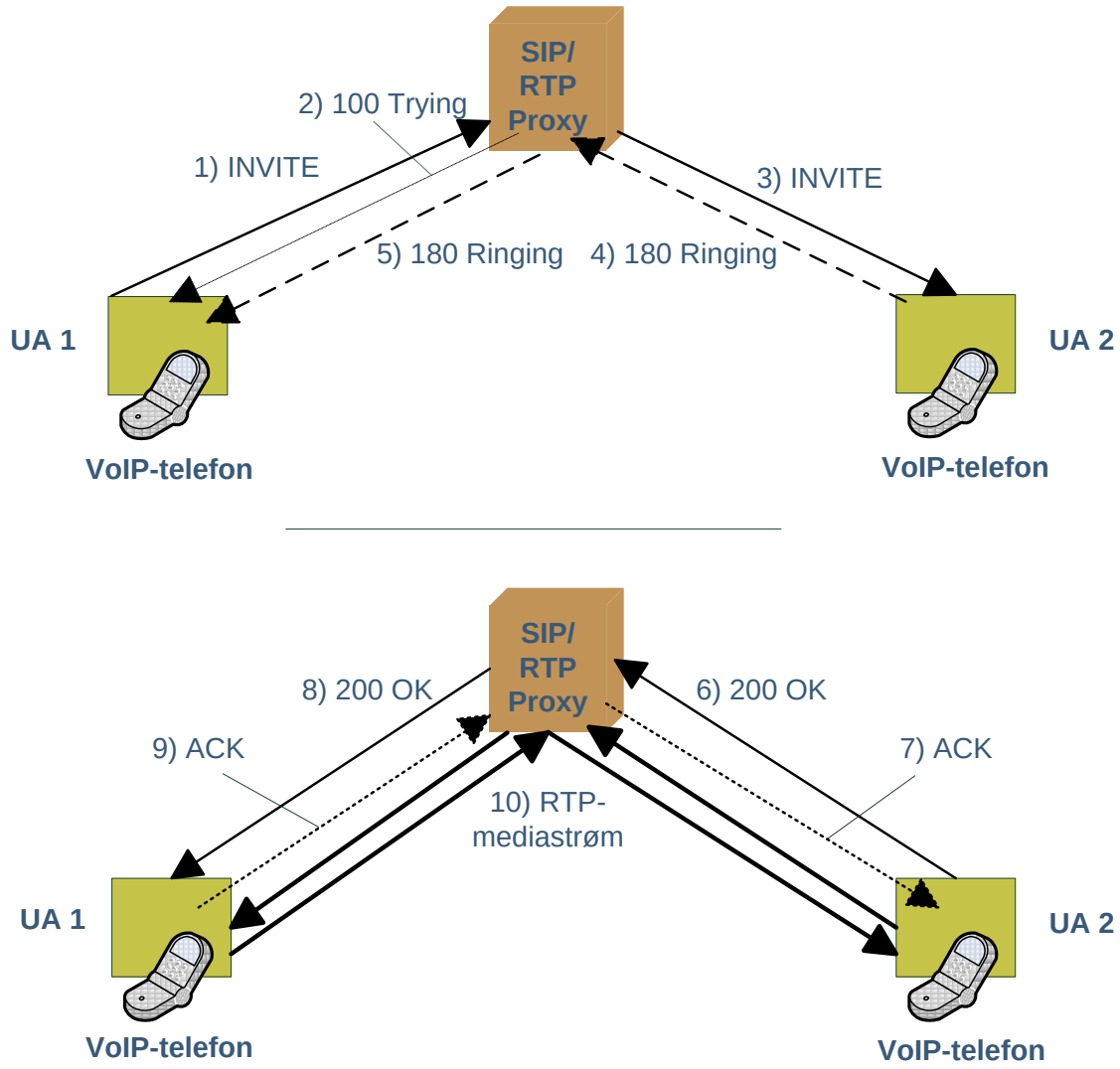
- UA1 sender en INVITE-forespørsel til UA2. Denne meldingen går først til SIP-proxy 1. INVITE-forespørselen adresserer UA2 ved hjelp av en SIP URI, og inneholder i tillegg SDP-informasjon om mediatype, mediaformat, transportprotokoll, IP-adresse og portnummer for å motta media hos UA1.
- SIP-proxy 1 ber UA1 om å autentisere seg ved å sende responsen ”407 Proxy Authentication Required” til UA1, som autentiserer seg ved å sende en ny INVITE-forespørsel
- Etter mottak av INVITE-forespørselen sender SIP-proxy 1 en ”100 Trying” respons tilbake til UA1 for å informere om at meldingen er mottatt og prosesseres. Hvis SIP-proxy 1 autentiserer UA1 vil den sende INVITE-forespørselen videre til SIP-proxy 2.
- SIP-proxy 2 sender en ”100 Trying” respons tilbake til SIP-proxy 1. SIP-proxy 2 slår opp URI i sin lokale database og finner IP-adressen til UA2. Når SIP-proxy 2 har sendt INVITE-forespørsel til UA2 svarer UA2 med å sende en ”180 Ringing” respons tilbake til UA1 via proxyene. Dette for å signalisere at telefonen til UA2 ringer.
- UA2 må gi beskjed om at den aksepterer forespørselen/samtalen (tar av røret) ved å sende en ”200 OK” responsmelding til UA1. Når ”200 OK” er en respons på en INVITE-forespørsel, vil den inneholde SDP-informasjon om mediakapabilitetene til UA2. Det vil si hvordan UA2 ønsker å sette opp samtalen og hvor UA2 ønsker å motta data.
- UA1 kjenner nå til UA2 sin lokasjon og sender en ACK direkte til UA2. Slik at UA2 vet at UA1 er klar til å sende og motta data.
- UA1 og UA2 setter så opp en direkte mediaforbindelse mellom seg hvor selve sesjonsinnholdet (tale) blir transportert.
- RTP-sesjonen avsluttes ved at en av partene, i dette tilfellet UA2, sender en BYE-forespørsel og den andre part responderer med ”200 OK”. BYE-forespørselen og ”200 OK” responsen vil gå direkte mellom UA1 og UA2 når RTP-trafikk går peer-til-peer.



Figur 3.5 Oppsett og terminering av RTP-sesjon peer-to-peer mellom UA i to domener

I Figur 3.5 blir RTP-sesjonen satt opp direkte mellom UA1 og UA2, men det er mer vanlig å sende RTP-data via RTP-proxy. Da vil UA2 og UA1 ikke få kjennskap til hverandres IP-adresser.

Figur 3.6 viser forenklet SIP-meldingsutveksling for oppsett av en RTP-sesjon som går via RTP-proxy. Utstyret er det samme i øverste og nederste del av figuren, men figuren er delt for lettere å få frem de forskjellige stegene (pilene) i meldingsutvekslingen. Siden mediasesjonen nå går via RTP-proxy må SIP-proxy sende en ACK-melding til UA2 når den mottar en ”200 OK” melding fra UA2. SIP-proxy sender så ”200 OK” meldingen videre til UA1 og mottar en ACK melding fra UA1. UA1 og UA2 har nå en mediaforbindelse via RTP-proxy.



Figur 3.6 SIP-anropsoppsett for en RTP-sesjon via RTP-proxy [5]

Det er problematisk at SIP-standarden finnes i mange versjoner og at det kommer stadig forslag til nye endringer via ”Request For Changes” (RFC) til IETF. Eksempler med forslag til endringer skjer innenfor: ”redirection”, ”proxying”, ”security”, ”registration”, ”mobility” og så videre. I mangel av en entydig standardiseringsprosess har dette ført til at SIP-nettverk for det meste blir bygget med utstyr fra samme leverandør.

Telekom-industrien med ETSI 3GPP og ETSI TISPAN arbeidsgrupper har forbedret situasjonen ved å definere en SIP-profil til bruk i ”IP Multimedia Subsystem” (IMS) nettverk. IMS er en kjernenettverksarkitektur som gjør det mulig for mobiloperatører å tilby brukere IP-baserte multimediatjenester som for eksempel VoIP. Dette har gjort SIP-protokollen mer entydig og dermed bidratt til å gjøre SIP IMS profilen interoperabel [7]. I dag overføres tale i mobilnett via 2G (GSM) og 3G (UMTS) selv om 4G (LTE) allerede er i bruk mange steder. 4G er et rent pakkesvitsjet nettverk. Det er klart at tale i LTE, ”Voice over LTE” (VoLTE), vil benytte SIP-profilen som er tatt frem for IMS. Årsakene til at LTE ikke allerede benyttes for tale er blant annet at LTE-terminaler først er blitt tilgjengelig i den senere tid, og at noen operatører velger å

vente med å lansere VoLTE til de har god nok LTE-dekning for å kunne gi god talekvalitet. Operatører kan også velge å støtte handover mellom LTE og UMTS, forutsatt at terminalen støtter begge nettverk, og vil da kunne lansere VoLTE tidligere. VoLTE vil sannsynligvis bli lansert i løpet av 2012. Det er usikkert om VoLTE eller 3dje parts aktører med andre VoIP-løsninger, vil dominere markedet for tale over LTE. Det finnes allerede i dag applikasjoner som kan lastes ned fra 3dje parts aktører med VoIP-telefoni som bruker bredbåndsforbindelsen til en 3G-smarttelefon.

3.2 Real-time Transport Protocol (RTP)

”Real-time Transport Protocol” (RTP) ble spesifisert av ”Internet Engineering Task Force” (IETF) [12]. RTP transporterer datapakker i sanntid og brukes både for tale og video. RTP bruker transportprotokollen UDP da sanntidslevering er viktigere enn pålitelighet. UDP gir bare best-effort-levering og garanterer ikke levering av pakker. RTP er hva som kan kalles en ”kontainerprotokoll”, hvis egenskap er å transportere ulike mediastrømmer (for eksempel tale) kodet med ulike algoritmer. Talekodingen av VoIP-dataene bestemmer kvaliteten og toleransen for feil i mediastrømmen. Forskjellige feilkorreksjonsalgoritmer kan til en viss grad rette opp problemet med tap av pakker.

RTP-pakken består av header og nyttelast, hvor header inneholder blant annet informasjon om:

- talekoder brukt i pakken (som sier noe om hva som sendes)
- pakkens sekvensnummer
- tidsstempel
- synkroniseringskilde (SSRC)
- bidragskilder (CSRC)

SSRC identifiserer kilden (utstyrsenhet) til en RTP-mediastrøm. Mottaker arrangerer pakkene i riktig rekkefølge ved å bruke sekvensnummer, tidsstempel og synkroniseringskilden. Hvis en RTP-mediastrøm er et samlet resultat fra flere RTP-mediastrømmer, vil CSRC gi informasjon om hver bidragsyter (utstyrsenhet). Dette vil for eksempel være tilfelle ved en talekonferanse med flere bidragsytere.

”Real-time Transport Control Protocol” (RTCP) blir ofte brukt sammen med RTP for å sende kontrollmeldinger til deltagere av en RTP-sesjon. RTCP gir blant annet informasjon om:

- forsinkelse
- jitter
- gjennomsnittlig pakketap
- synkroniseringskilde (SSRC)
- bidragskilder (CSRC)
- deltagers navn og mailadresse [7]
- mapping av deltager mot kilde for RTP-mediastrøm

RTP og RTCP bidrar til at mottakere kan takle jitter og andre problemer ved passende bufring og sekvensering. I tillegg vil mer informasjon om nettverket bidra til at en kan gjøre forebyggende tiltak som å innføre mer redundans eller lavere koderate etc. Når RTP og RTCP brukes sammen over UDP vil RTP bli tilegnet en liketalls UDP-port, mens RTCP blir tilegnet den neste oddetalls UDP-porten. Det vil si at hver deltager av en RTP-sesjon bruker minst to UDP-porter hver, i området 1024 til 65535 [3;5-7].

4 Sikkerhet i SIP og RTP basert VoIP og multimediasesjoner

Sikkerhet i VoIP omhandler hvordan en kan sikre VoIP-trafikk med hensyn til sikkerhetsegenskapene; konfidensialitet, integritet og tilgjengelighet. Dette kapittelet ser på sikkerhetstrusler og angrep i SIP- og RTP-basert VoIP, hvilke sikkerhetsegenskaper som er nødvendig for å motvirke disse og sikkerhetsmekanismer som brukes i praksis for å oppnå disse sikkerhetsegenskapene. Tabell 4.1 viser noen eksempler på mulige trusler og sammenhengen mellom trusler, sikkerhetstjenester og sikkerhetsmekanismer. For hver trussel er det en sikkerhetstjeneste som kan imøtekomme trusselen, hvor sikkerhetstjenesten realiseres ved hjelp av en sikkerhetsmekanisme. Denne rapporten fokuserer på tilgjengelige sikkerhetsmekanismer i protokollene SIP og RTP. Sikkerhetsaspektene ved SIP og RTP vil også være dekkende for video og andre typer multimediasesjoner som bruker SIP- og RTP-protokollene.

Trusler/angrep	Sikkerhetstjenester	Sikkerhetsmekanismer
Misbruk av identitet (Identity Fraud)	Autentisering Tilgangskontroll	Autentiseringsmekanismer Aksess lister, SIP-peering
“Spam Through Internet Telephony” (SPIT)	Autentisering Tilgangskontroll	Autentiseringsmekanismer “White- and blacklists”
Avlytting av signalering	Kryptering	Krypteringsmekanisme

Tabell 4.1 Sammenheng mellom trusler, sikkerhetstjenester og sikkerhetsmekanismer i VoIP [13]

Når det gjelder de proprietære protokollene for realisering av VoIP og andre multimediasesjoner, som Skype og Google Talk, så har disse gode sikkerhetsløsninger med god kryptering.

VoIP-trafikken er lettest tilgjengelig og mest sårbar på hoppet mellom endebruker (UA) og SIP- og media-server. Her kan det være mangelfulle sikkerhetsmekanismer og/eller benyttelse av delte aksess-nett (som for eksempel WLAN). Trafikken mellom VoIP-leverandørene, eller internt i en leverandørs nett, er vanskelig for en angriper å nå uten først å få kontroll over node(r) hvor VoIP-trafikken strømmer gjennom. En leverandør med en utro tjener, som har de rette tilgangene, vil ikke ha denne begrensningen.

4.1 Sikkerhetstrusler og angrep i SIP og RTP basert VoIP

En sårbarhet kan defineres som en feil eller svakhet i protokolldesign, implementasjon eller installasjon av VoIP-systemer. En trussel er en potensiell utnyttelse av en sårbarhet. "The VoIP Security Alliance" (VOIPSA) har kommet frem til en klassifisering av ulike sikkerhetstrusler mot VoIP. Denne klassifiseringen utgjør et rammeverk for å kategorisere sikkerhetstrusler mot VoIP-installasjoner, -tjenester og -brukere. I dette rammeverket er det definert fem kategorier med trusler [14]:

- Sosiale trusler ("Social threats") – inkluderer identitetstyveri samt feilaktig presentasjon av identitet og innhold.
- Avlytting ("Eavesdropping") – avlytting av signallering (SIP) eller multimediastrømmen (RTP) uten å endre innhold.
- Avlytting og endring ("Interception and modification") – avlytting og endring av signallering og/eller multimediainnhold
- Misbruk av tjenester ("Service abuse") – beskriver trusler som favner feilaktig bruk av en VoIP-leverandørs tjenester. Trusler inkluderer blant annet å omgå VoIP-leverandørs autentiseringsmekanisme og registrering, se avsnitt 4.1.1. Motivasjonen for å gjennomføre disse truslene er som regel økonomisk vinning.
- Avbrudd av tjeneste ("Interruption of service") – Angrep i denne kategorien kan innvirke på et hvilket som helst nettverkelement som er en del av VoIP-tjenesten, inkludert rutere, DNS-server, SIP-proxy, "Session Border Controllers" (SBC) med mer. Kategorien inkluderer blant annet "Denial of Service" (DoS) trusler. "Spam Through Internet Telephony" (SPIT) kan tillegges denne kategorien hvis en utvider den også til å dekke irritasjon ("Annoyance").

En sikkerhetstrussel som faller inn i en eller flere av VOIPSA-kategoriene vil krenke en eller flere av sikkerhetsegenskapene; konfidensialitet, integritet og tilgjengelighet. Hvis for eksempel en angriper kan lese og modifisere pakkene i et nettverk vil det krenke både systemets konfidensialitet og integritet.

VoIP-industriaktører i Norge er med hensyn til sikkerhet mest bekymret for identitetsbedrageri, misbruk av tjenester, og økonomisk tap på grunn av misbruk av utgående samtaler. Disse truslene kan kategoriseres etter VOIPSA sitt rammeverk som "Social threats", "Interception and modification" og "Service abuse" [15].

Et sikkerhetsangrep er en implementert trussel, mer presist definert som et forsettlig angrep på systemet utledet fra en trussel. I litteraturen er sikkerhetsangrep klassifisert som passive og aktive:

- Passive angrep – avlytter kommunikasjon og tilegner seg informasjon ved informasjonslekkasje. Det er definert to typer passive angrep; 1) tilgang til mediastrømmen (RTP-nyttelast) ved å lytte seg inn på en VoIP-samtale og 2) trafikkanalyse. Trafikkanalyse er analyse av trafikkmønsteret som går over en link hvor

trafikk, volum og så videre kan gi verdifull informasjon. For eksempel vil en økning i trafikk kunne si noe om aktivitetsnivået til de som analyseres. Trafikkanalyse bruker blant annet informasjon om lengde på kommunikasjonsstrøm, når den ble sendt og de involverte UA, til å observere kommunikasjonsmønsteret til en eller flere UA. Tilgjengelig informasjon fra SIP-headere, SDP-nyttelast og RTP-header vil kunne brukes i en trafikkanalyse. Denne informasjonen kan videre brukes til å utføre forskjellige typer aktive angrep.

- Aktive angrep – involverer modifisering av data eller injeksjon av falske data. Aktive angrep kan utføres mot både SIP- og RTP-meldingene [15].

Et vellykket angrep er i virkeligheten ofte en kombinasjon av passive og aktive angrep. De største truslene mot SIP i praksis er informasjonslekkasje ved avlytting, endring av signalleringsdata uten at det oppdages, misbruk av identitet og "Distributed DoS" (DDoS) angrep. De største truslene mot RTP i praksis er informasjonslekkasje ved avlytting av RTP-meldingsinnhold. Se Tabell 4.2 for eksempel på sikkerhetsmekanismer som kan brukes for å sikre SIP- og RTP-protokollen. Sikkerhetsmekanismene blir beskrevet nærmere i avsnitt 4.2 og 4.3 og DDoS i avsnitt 4.1.5.

Trusler/angrep	Sikkerhetstjeneste for	Sikkerhetsmekanismer
Identitetsforfalskning	Autentisering	Secure SIP (TLS), S/MIME
Endring av signalleringsdata	Integritetsbeskyttelse	Secure SIP (TLS), S/MIME
Informasjonslekkasje av SIP-headere	Konfidensialitetsbeskyttelse	Secure SIP (TLS)
Informasjonslekkasje av SDP-nyttelast	Konfidensialitetsbeskyttelse	Secure SIP (TLS), S/MIME
DDoS-angrep	Tilgjengelighet	Vanskelig å beskytte seg mot
Informasjonslekkasje av mediastrøm (RTP-nyttelast)	Konfidensialitetsbeskyttelse	SRTP
Informasjonslekkasje av RTP-headere	Konfidensialitetsbeskyttelse	Kan beskyttes ved bruk av IPSec, VPN og DTLS
Endring av RTP-header og RTP-nyttelast	Integritetsbeskyttelse	SRTP
Forfalskning av pakker	Autentisering	SRTP

Tabell 4.2 Sikkerhetsmekanismer som kan brukes for å sikre SIP og RTP-protokollen

Under følger noen eksempler på klassiske trusler/angrep som demonstrerer behovet for spesifikke sikkerhetstjenester. Disse sikkerhetstjenestene kan potensielt forhindre hele kategorier av trusler. Angrepene er valgt for å illustrere viktigheten ved sikkerhetstjenester i SIP. Angripers motivasjon kan være å stjele tjenester, tilegne seg informasjon ved avlytting og/eller avbryte sesjoner [9].
En forutsetning som ligger til grunn for å muliggjøre disse angrepene er at angriper må ha full kontroll over nettverkstrafikken til og fra en VoIP-klient, slik at angriper kan avlytte og modifisere pakker på nettverket.

4.1.1 Kapring av registrering (”Registration Hijacking”)

Når telefonen (UA) blir slått på, og ved regelmessige intervaller (3-60 min), må UA registrere seg mot en lokal registreringsserver. Denne registreringsserveren forventer og krever som regel autentisering for hver registreringsforespørsel fra UA. Registrering er nødvendig for at lokal VoIP-server skal vite hvilken IP-adresse/vertsnavn (kontaktadresse) UA kan nås på ved innkomne anrop. Et ”registration hijacking” angrep vet å utnytte en sårbarhet i denne transaksjonen. Angriper gir seg ut for å være autorisert bruker for således å forlede registreringsserver til å registrere angripers kontaktadresse (IP-adresse/vertsnavn til telefonen) som UA sin kontaktadresse. Konsekvensene av dette er at VoIP-serveren sender alle forespørsler til angriper istedenfor til autorisert bruker. Se avsnitt 4.2.1 for mer informasjon om kapring av registrering.

4.1.2 Utgi seg for å være server (”Impersonating a Server”)

Angriper kan utgi seg for å være en legitim SIP-server (proxy, registrar, redirect) og fange opp forespørsler fra UA og manipulere disse. Redirect vil si når server gir beskjed til sender at den må kontakte et annet sett av URIer.

Eksempel: UA1 kontakter sin lokale VoIP-server sip.oslo.no, hvor den først registrerer seg og prøver å ringe UA2. Angriper har gitt seg ut for å være sip.oslo.no (ulike angrep for å få til det) og dermed vil all kommunikasjon fra UA1 gå gjennom angripers VoIP-server. Angriper har således tilgang til all kommunikasjon som kommer fra UA1. Angriper er ”man-in-the-middle” og kan:

- endre signallering og data (integritetsbrudd)
- avlytte signallering og data (konfidensialitetsbrudd)

For å motvirke disse truslene må bruker kunne autentisere VoIP-serveren. Autentiseringen som brukes i SIP er DAA. DAA autentiserer bare klient (”en-veis autentisering”), så disse truslene er høyst reelle i SIP i dag. Se avsnitt 4.2.1 for mer informasjon om DAA.

4.1.3 Terminering av sesjoner (”Tearing down sessions”)

En etablert RTP-sesjon kan endres ved å sende SIP-forespørselen ”BYE” eller ”re-INVITE” til deltagerne av sesjonen. Hvis deltagere i RTP-sesjonen ikke kan validere hvem disse forespørslene kommer fra, åpner det blant annet for at angriper kan ta ned samtaler.

En angriper kan tilegne seg informasjon ved å lytte til signalleringen for en sesjon. Angriper kan så bruke denne informasjonen til å sende en falsk SIP BYE-forespørsel til deltagerne av sesjonen. Denne forespørselen må være konstruert til å se ut som den kommer fra en sesjonsdeltager. Konsekvensene av dette er at RTP-sesjonen blir avsluttet før deltagerne ønsker det. Det er også mulig at angriper kan sende en falsk re-INVITE for å endre på sesjonen slik at sikkerheten svekkes (”down grade”-angrep) eller for å omdirigere trafikken i en konferansesamtale i et avlyttingsforsøk.

Disse truslene kan motvirkes ved krav om autentisering av den som sender BYE-forespørselen og re-INVITE-forespørselen. Hvis signalleringen er kryptert, har ikke angriper mulighet til å tilegne seg sesjonsinformasjon og kan således ikke sende en falsk BYE- eller re-INVITE-forespørsel.

Det er ikke autentisering av sender av BYE- eller re-INVITE-forespørsler i dag, se avsnitt 4.2.1. Det er heller ikke vanlig å kryptere SIP-signalleringen, noe som gjør trusselen beskrevet i dette avsnittet høyst reell.

4.1.4 Misbruk av tjenester (”Service Abuse”)

Kapring av en SIP-sesjon (”call hijack”) er vist og forklart i [15;16]. Angriper tilegner seg nødvendig informasjon om deltagerene under en SIP-INVITE handshake. Deretter kaster angriper ut sender og mottaker ved å sende falske SIP-CANCEL meldinger til begge parter. Angriper manipulerer så resten av SIP-INVITE handshaken til å sette opp sin egen samtale. Siden dette angrepet er vanskelig å oppdage av SIP-proxyserver vil samtalen bli registrert som en vanlig samtale. Den opprinnelige eier av samtalen må betale for samtalen til angriper, og er utsatt for tjenestemisbruk. Det kreves ikke autentisering av sender av SIP-CANCEL forespørsel i dag, men det er allikevel ganske omfattende forutsetninger som ligger til grunn for at dette blir et reelt angrep.

4.1.5 Tjenestenektangrep (“Denial of Service”)

En av de mer ”tradisjonelle” og vanligste angrepene på Internet i dag er hva som kalles tjenestenektangrep, mer kjent som ”Denial of Service” eller bare ”DoS”. DoS-angrep er enkle å utføre og vanskelig å beskytte seg mot. Den vanligste formen for DoS-angrep er å sende et stort antall meldinger som overbelaster tjenesten. Konsekvensen er at tjenesten ikke har ressurser til å besvare legitime forespørsler, eller i verste fall går ned. DoS-angrep er gjerne utført distribuert og koordinert av flere angrepsmaskiner for å forsterke angrepet mot en tjeneste. Disse kalles således ”Distributed DoS”-angrep.

DoS-angrep mot VoIP-installasjoner kan slå ut telefonitjenesten, og dermed gjøre det vanskelig eller umulig for brukere å sette opp sesjoner (ringe). De kan også medføre forringet talekvalitet. DoS-angrep mot VoIP er en høyst reell trussel, og angrep skjer stadig mot VoIP-installasjoner i dag [6].

4.2 Sikkerhet i SIP-signallering

Dette er de tre hovedgrunnene til at SIP er vanskelig å sikre:

1. SIP er designet med hensyn til funksjonalitet og ikke sikkerhet.
2. SIP er avhengig av mellomledd (for eksempel SIP-proxy) som ofte trenger å lese og modifisere SIP-trafikken.
3. SIP-standarden har økt i omfang og kompleksitet de senere årene som gjør den ytterligere sårbar. Nye tillegg til SIP-protokollen innfører også nye sikkerhetsutfordringer.

Under følger en gjennomgang av forskjellige sikkerhetsmekanismer som kan brukes for å sikre SIP-signallering; DAA-autentisering, Secure SIP og S/MIME. I tillegg nevnes ny forskning på forbedring av SIP-autentisering.

4.2.1 ”Digest Access Authentication” (DAA) autentisering

Den mest brukte autentiseringsmetoden i SIP i dag er ”Digest Access Authentication” (DAA). Standarden krever at SIP skal støtte DAA, men ikke at SIP må bruke DAA. DAA har stor utbredelse siden støtte for den er påkrevd i standarden, den er enkel å implementere og bruke/konfigurere. Den benyttes kun innenfor ett domene, og tilbyr ikke ende-til-ende autentisering på tvers av domener. Som tidligere nevnt er et domene en fysisk eller logisk avgrensning, gjerne ett bestemt nettverk eller en bedrift. Denne autentiseringsmetoden brukes primært i SIP-REGISTER og SIP-INVITE handshakes. DAA gir enveis-autentisering og replaybeskyttelse. Alle SIP-meldinger blir sent i klartekst [17;18].

DAA benytter en MD5-hash funksjon hvor server verifiserer klienten (VoIP-bruker) sitt passord. Server generer en tilfeldig engangsverdi (”nonce”) og sender denne til klienten (UA). Bruken av nonce-verdi er ment å forhindre replay-angrep. Server kan enten være en ”proxy server”, ”redirect server” eller ”registrar”. ”Redirect server” er en logisk enhet som gir beskjed til avsender at den må kontakte mottaker på en annen URI. Klienten beregner en MD5-hash over ”nonce”, brukernavn, passord og noen andre SIP-header verdier som utgjør en DAA-responsverdi. Når server mottar DAA-responsverdien fra klient, utfører server samme beregning og sammenligner de to verdiene. Hvis de to er identiske er klienten (UA) autentisert. Svakheter med DAA:

- Bare autentisering av bruker (enveis-autentisering)
- Få SIP-header verdier som blir integritetsbeskyttet
- Ingen konfidensialitetsbeskyttelse av SIP-meldingen (bare passordet blir beskyttet)
- DAA bruker MD5 som har fått påvist flere svakheter de siste årene

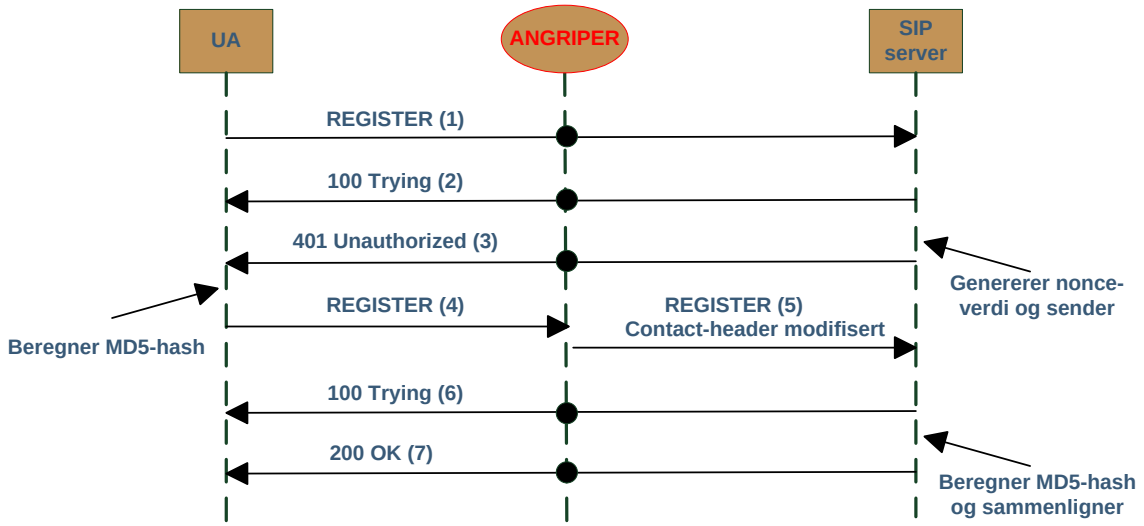
Det at DAA bare gir enveis-autentisering av bruker kan som tidligere nevnt i avsnitt 4.1.2 også føre til MitM-angrep, ved at angriper kan gi seg ut for å være en SIP-proxyserver. Da vil angriper få tilgang til all signallerings- og multimediatrafikk som går mellom to UAer og kan manipulere denne trafikken.

Siden SIP-meldingene blir sendt i klartekst er det mulig for angriper å tilegne seg relevant informasjon om sender og mottaker. Angriper kan bruke denne informasjonen til å sende falske SIP-re-INVITE, -CANCEL og -BYE forespørsler, siden disse forespørslene som regel ikke krever autentisering av avsender. Dette kan føre til angrep som beskrevet i avsnitt 4.1.3 og 4.1.4. Angriper vil også kunne tilegne seg informasjon om sender, mottaker, type samtale og varighet for bruk i trafikkanalyse. Dette kan være interessant hvis mediastrømmen er kryptert eller angripers motivasjon først og fremst er å følge med på trafikkmønsteret til VoIP-brukerne.

Det er en stor svakhet ved DAA-autentiseringen at en angriper kan utføre et man-in-the-middel (MitM) angrep på SIP-registreringen, uten å kjenne til den delte hemmeligheten (passordet)

mellom klient og server. Dette utgjør en betydelig sårbarhet i autentiseringen til SIP og er forårsaket av at for få SIP-headerverdier beskyttes av DAA. I registreringsprosessen vil klienten (UA) etter en vellykket DAA-autentisering få registrert og knyttet sin identitet (SIP URI) til et vertsnavn eller IP-adresse hos en lokasjonsserver (Registrar). Re-registreringer utføres som sagt som regel hvert 3-60 min avhengig av konfigurasjon. I UA sin SIP-REGISTER melding inneholder SIP-headeren ”Contact”, informasjon om hvilke vertsnavn/IP-adresse SIP-brukeren benytter. Denne informasjonen er sendt i klartekst, er ikke beskyttet av DAA, og er sårbar for MitM-angrep. En angriper kan endre vertsnavn eller IP-adresse til klienten i registreringsmeldingen i sanntid ved hjelp av for eksempel nettverksverktøyet ”NetSED”, uten at hverken klienten eller Registrar oppdager det. Registrar vil registrere angripers IP-adresse som kontaktpunkt. Når så en SIP-server gjør et oppslag i lokasjonsserveren (Registrar) for å finne hvilket vertsnavn/IP-adresse UA kan nås på, vil sesjonen rutes til et vertsnavn eller IP-adresse som er kontrollert av angriper. Det er ingen mekanismer for å detektere denne type angrep og UA vil være uvitende om at den ikke kan nås. Konsekvensene av angrepet skissert er at angriper kan ta over identiteten til en UA uten at server og UAen selv er klar over det. Angriper må ha tilgang til nettverkstrafikken, det vil si at angriper må være på ”innsiden” av lokal-domenet/bedriftens nett [18].

Figur 4.1 viser MitM-angrep på SIP-registreringen. UA sender en registreringsforespørsel (1) til SIP-server og får en respons (2). SIP-server generer en nonce-verdi og sender den til UA med et krav om autentisering (3). UA beregner en MD5-hash og sender denne i en ny registreringsforespørsel (4). Angriper gir seg ut for å være SIP-server (MitM) og modifiserer vertsnavn/IP-adresse i Contact-header i registreringsforespørselen (5) som han sender videre til SIP-server. SIP-server har autentisert UA (7) og har registrert angripers kontaktadresse istedenfor brukerens [18].



Figur 4.1 Angriper modifiserer Contact-header verdi i SIP-registreringsmelding [18]

Det er også en svakhet at DAA benytter MD5 som er relativt enkel å knekke ved ”brute-force”, som vil si at angriper gjetter seg frem til brukers passord ved å prøve og feile. Dette forutsetter at

angriper har hatt mulighet til å avlytte en sesjon tidligere, slik at angriper har tilegnet seg alle verdier som inngår i DAA-hashen foruten passordet. Siden angriper også sitter på DAA-resultatet, vet han når han har gjettet riktig. Hvis angriper får tak i brukers passord kan han misbruke brukers identitet og misbruke tjenester, ved blant annet å sende falske SIP-REGISTER eller SIP-INVITE meldinger på vegne av bruker.

Det er i tillegg påvist svakheter i MD5-hash funksjonen. MD5 er blant annet ikke lenger kollisjonssikker og kan bli utsatt for kollisjonsangrep, hvor to forskjellige inputverdier produserer samme MD5-hash. Det vil si at to brukere med to forskjellige passord kan få samme DAA-resultat, som vil si at server ikke kan se forskjell på dem. Denne svakheten er ikke kjent for å kunne utnyttes til å avsløre brukers passord. Likevel er det anbefalt å bruke en sterkere hashfunksjon som for eksempel SHA-2 [18;19].

DAA bør byttes ut med en bedre autentiseringsmetode. Lars Strand og Wolfgang Leister fremlegger teoretiske forslag til en sikrere autentisering i SIP i [18]. På kort sikt foreslår de en modifisert ”Password Authenticated Key Exchange” (PAKE) autentisering som erstatning for DAA. Fordeler med den modifiserte PAKE-autentiseringen:

- Tosidig autentisering av både bruker og SIP-server
- Gjenbruk av det delte passordet som gjør det enklere å erstatte DAA
- Sterk beskyttelse av passord som gjør brute-force angrep vanskeligere å gjennomføre

På sikt er det ønskelig med en mer fleksibel autentiseringsmekanisme, da forskjellige VoIP-installasjoner har forskjellige sikkerhetskrav som kan behøve forskjellige sikkerhetstjenester. Lars Strand og Wolfgang Leister foreslår derfor på sikt innføring av sikkerhetslaget ”Simple Authentication and Security Layer” (SASL), som tilbyr et grensesnitt for autentisering og forhandling av autentiseringsmekanismer. Fordeler med SASL:

- fleksibel og adaptiv til nye sikkerhetskrav og fremtidige endringer ved å gi tilgang til flere underliggende autentiseringsmekanismer
- minimale endringer av SIP-standarden er nødvendig
- moden, stabil og velprøvd i industrien

Det gjenstår å se om disse nye forslagene blir en del av SIP-standarden, implementert og tatt i bruk i industrien.

4.2.2 Autentisering og integritetsbeskyttelse ved bruk av S/MIME

SIP støtter også bruk av autentiseringsmekanismen ”Secure Multipurpose Internet Mail Extensions” (S/MIME), som har til hensikt å oppnå ende-til-ende autentisering mellom to UAer [9].

Ved S/MIME så kapsles den originale SIP-meldingen inn i en MIME-body som signeres og /eller krypteres. Denne blir så fraktet som payload i en ny ”ytre” SIP-melding. Mottaker sjekker signaturer/dekrypterer og pakker ut MIME-body for å få frem original SIP-melding [9].

Autentisering i S/MIME blir utført ved hjelp av digitale sertifikater, vanligvis X.509 sertifikater. Det er derfor behov for en ”Public Key Infrastructure” (PKI) for distribusjon av sertifikater til hver deltager fra ”Certificate Authority” (CA), som er en betrodd 3.-djepart. Da det ikke finnes en samlet global CA, må UA støtte flere rotsertifikater hvis den kommuniserer med en UA som bruker en annen CA.

Eksempel: Telenor og Netcom har blitt enige om å støtte CA X. Leverandøren Orange vil bruke CA Y. Dermed må Telenor- og Netcom-kunder også stole på CA Y for å kunne validere Orange sine kunder.

PKI er både vanskelig og kostbart å implementere. I følge den årlige ”SCI Computer Crime and Security Survey” (2011) har bare 35 % av selskapene i undersøkelsen investert og tatt i bruk en PKI. Selskapene i undersøkelsen omfatter amerikanske aksjeselskap, offentlige organer, finansinstitusjoner, læresteder, medisinske institusjoner og andre organisasjoner. Det gjenstår fremdeles å finne en løsning på PKI i telefoni som tilbyr skalerbarhet i forhold til bruk i mange forskjellige telefoner og kommunikasjonsenheter. I tillegg vil sertifikathåndtering som annullering og fornyelse av sertifikater komplisere bruken av sertifikater [9;15;17].

Det har vært begrenset støtte for S/MIME i industrien. På ”International SIP Interoperability Test Conference 28” (SIPit 28) støttet ingen av de 40 ulike SIP-implementasjonene S/MIME [15].

4.2.3 Secure SIP (SIPS) over TLS

“Secure SIP” (SIPS) er en sikkerhetsmekanisme definert i SIP RFC 3261 for å sende SIP-meldinger over en ”Transport Layer Security” (TLS) kryptert kanal. SIPS tilbyr tosidig autentisering, integritets- og konfidensialitetsbeskyttelse av SIP-meldinger på hopp-per-hopp basis. Hver SIP-server må ha tilgang til en del informasjon i SIP-meldingen slik at TLS-forbindelsen må termineres og initieres for hvert hopp. TLS kan brukes mellom domener. Avsender setter opp en TLS-forbindelse til nærmeste SIP-proxyserver, men har ingen garanti for at TLS blir benyttet hele veien frem til mottaker. Avsender får heller ingen informasjon om TLS er blitt brukt hele veien frem til mottaker. SIPS vil beskytte mot identitetsforfalskning og informasjonslekkasje av SIP-data. I tillegg til at endring av signalleringsdata vil bli oppdaget. SIPS vil dermed beskytte mot angrepene beskrevet i avsnittene 4.1.1 til 4.1.4 [9;18].

Secure SIP setter opp en kryptert TLS-forbindelse mellom UA og SIP-server ved hjelp av ”Public Key Encryption”, hvor hver part har en offentlig og en privat nøkkel. For at UA og SIP-server skal være sikre på at de snakker med ”riktig” nøkkel (identitet) så vil meldingene inneholde et sertifikat som er utstedt fra en CA. Den offentlige og private nøkkelen er en del av sertifikatet, og hver part har fått sine nøkler ”signert” av en CA. UA og SIP-server autentiserer hverandre ved å validere hverandres sertifikater. Hvis sertifikatene er signert av en CA begge stoler på, vil de stole

på hverandre. Denne autentiseringen kan være enveis eller toveis. Når UA og SIP-server har validert hverandres sertifikater, vil sertifikatene (nøklene) bli brukt til å forhandle frem TLS-forbindelsen. Når TLS-forbindelsen er etablert settes det opp en vanlig SIP-forbindelse. SIP-REGISTER med DAA-autentisering gjøres nå gjennom en kryptert TLS-forbindelse. Siden TLS benytter seg av x509-sertifikater og PKI har Secure SIP de samme utfordringene som S/MIME diskutert tidligere i 4.2.2.

SIPS bruker standardiserte kryptoalgoritmer som er anbefalt av blant annet NSM. Den anbefalte krypteringsalgoritmen i SIPS er AES i ”Cipher Block Chaining” (CBC) modus med bruk av en 128-bit nøkkel. For integritetsbeskyttelse vil hash-funksjonen SHA-1 bli brukt for utregning av ”Message Authentication Code” (MAC). SHA-1 er ikke lenger kollisjonssikker, men kollisjonssikkerhet er ikke kritisk når SHA-1 blir brukt for utregning av MAC. [6;9;15;19].

Det kan være ressurskrevende for en SIP-server å opprettholde mange samtidige TLS-forbindelser, spesielt ved bruk av sterke kryptoalgoritmer. Dette kan føre til skaleringsproblemer. TLS er også utfordrende ved bruk av mobile enheter grunnet begrenset batteri og/eller CPU [9].

Secure SIP fordrer bruk av TCP, men ofte benytter VoIP-installasjoner i industrien seg av SIP over UDP. Det er derfor gjort standardiseringsarbeid i den senere tid for å støtte TLS over UDP i SIP, som har resultert i ”Datagram Transport Layer Security” (DTLS) [15].

”Secure SIP” har vunnet noe støtte i industrien og det ser ut som industrien er i ferd med å velge SIPS fremfor S/MIME, sannsynligvis fordi TLS er en utbredt industristandard. I praksis blir S/MIME og andre sikkerhetsmekanismer lite brukt, men SIPS kan bli mer fremtredene når flere leverandører implementerer dette samt løser utfordringene med tanke på PKI i telefoni. På ”International SIP Interoperability Test Conference 28” (SIPit 28) støttet 50% av de 40 ulike SIP-implementasjonene TLS med tosidig autentisering. I RFC 3261 anbefales det å ta i bruk TLS på SIP-servere [15].

4.2.4 Oppsummering sikkerhet SIP

Det som benyttes av sikkerhetsmekanismer for SIP-signallering i dag er hovedsakelig DAA-autentisering, som har vist seg svært sårbar. SIPS og S/MIME lider begge under utfordringene ved håndtering av X.509 sertifikater, i tillegg til kompleksiteten og kostnadene ved implementering av en PKI. Det gjenstår fremdeles å finne en løsning på PKI i telefoni som tilbyr skalerbarhet i forhold til bruk i mange forskjellige telefoner og kommunikasjonsenheter. SIPS har vunnet noe støtte i industrien og det ser ut som industrien er i ferd med å velge SIPS som felles sikkerhetsmekanisme, sannsynligvis fordi TLS allerede er en etablert industristandard.

Oppsummering av sikkerhetsdesign i SIP:

- SIP-meldinger blir sendt i klartekst. Informasjon om sender, avsender, type samtale og varighet er derfor tilgjengelig for angriper.

- DAA integritetsbeskytter for få SIP-headere. Dette åpner for MitM-angrep på SIP-REGISTER meldingen. Angriper kan modifisere SIP-registreringen til å registrere brukers SIP URI opp mot angripers eget vertsnavn eller IP-adresse. Innkomne anrop rutes dermed til angripers vertsnavn eller IP-adresse istedenfor til legitim bruker. Dette angrepet kan utføres ubemerket for både klient og server.
- Angrep på SIP-INVITE ”handshake”, ved hjelp av falske SIP-CANCEL forespørsler og noe annen modifikasjon av SIP-dialogen, resulterer i hijacking av hele samtalen. Det vil si at angriper tar over SIP-sesjonen og bruker blir fakturert for samtalen.
- Angriper kan sende falske SIP-BYE forespørsler til sender og mottaker og avslutte en samtale.
- DAA autentiserer bare bruker og det kan føre til MitM-angrep ved at angriper gir seg ut for å være SIP-proxyserver. Da vil angriper få tilgang til og mulighet til å endre alle signallerings- og brukerdata som går mellom to UAer.
- DAA er relativt enkel å knekke ved et brute-force-angrep. Ved et vellykket bruce-force-angrep vil angriper få tilgang til passordet til bruker og kan utnytte brukers identitet. Angriper kan blant annet sende falske SIP-REGISTER og SIP-INVITE forespørsler.

Fordeler ved S/MIME:

- S/MIME opererer på applikasjonslaget og kan bli brukt både med UDP og TCP
- Vil kunne gi ende-til-ende autentisering ved validering av sertifikat
- Vil kunne gi ende-til-ende integritets- og konfidensialitetsbeskyttelse
- Tilbyr stor fleksibilitet ved at den kan beskytte deler av SIP-meldingen

Begrensninger ved S/MIME:

- Krevende å implementere på grunn av dens kompleksitet og at den behøver en PKI-infrastruktur
- Har lite støtte i industrien og blir i praksis ikke brukt
- Skalering kan bli et problem på grunn av behovet for en PKI-infrastruktur

Hvis Secure SIP blir brukt i tillegg til DAA vil det beskytte mot sikkerhetstruslene nevnt over.

Fordeler ved bruk av SIPS:

- SIPS tilbyr tosidig autentisering, integritets- og konfidensialitetsbeskyttelse av SIP-meldinger på hopp-per-hopp basis
- SIPS-autentiseringen kommer i tillegg til DAA-autentiseringen
- SIPS kan integritets- og konfidensialitetsbeskytte forhandling av kryptografiske nøkler for beskyttelse av mediastrømmen ved bruk av SRTP
- SIPS benytter TLS, som er en utbredt, moden og grundig testet industristandard

Begrensninger ved bruk av SIPS:

- Gjenstår fremdeles å finne en løsning på PKI i telefoni som tilbyr skalerbarhet i forhold til bruk i mange forskjellige telefoner og kommunikasjonsenheter
- SIPS gir ingen garanti for ende-til-ende beskyttelse og bruker får heller ingen informasjon om den har fått ende-til-ende beskyttelse
- Hvis SIP-proxyserver er infiltrert av angriper vil angriper ha tilgang til signallering og SDP-nyttelast
- Ved bruk av TLS blir SIP begrenset til å bruke transportprotokollen TCP. Dette er et problem da SIP ofte blir sendt over UDP i virkelige industriscenarier. Det er gjort arbeid for at TLS kan bruke UDP, det vil si ”Datagram Transport Layer Security” (DTLS). Og det er vurdert tillegg til SIP-protokollen for å bruke SIP over DTSL

4.3 Sikkerhet i RTP

Standardprotokollen for beskyttelse av sanntid-media (tale og video) i multimedia-applikasjoner er ”Secure Real-time Transport Protocol” (SRTP), definert i [20]. Den gir autentisering, integritets- og konfidensialitetsbeskyttelse av mediastrømmen for å beskytte mot angrep som å tilegne seg informasjon ved avlytting, endring av innhold og forskjellige DoS-angrep. I tillegg til å beskytte RTP-pakker vil den også kunne beskytte RTCP-meldinger. Applikasjonen som implementerer SRTP må omdanne henholdsvis RTP- og RTCP-pakker til SRTP- og SRTCP-pakker før pakkene blir sendt over nettverket. Samme prosess blir reversert ved mottaker for å dekryptere SRTP- og SRTCP-pakkene tilbake til RTP- og RTCP-pakker.

Det er anbefalt som et minimum å bruke autentisering og integritetsbeskyttelse av RTP-pakker. Det optimale er å bruke både autentisering, integritetsbeskyttelse og kryptering av meldinger. SRTP blir ikke brukt som standard praksis for tale og video i dag. En grunn kan være at VoIP-leverandørene har vært sene med å implementere SRTP i sine produkter, og at det medfører økt teknisk kompleksitet og ekstra kostnader å ha slik funksjonalitet [6].

4.3.1 ”Secure Real Time Protocol” (SRTP)

Standard krypteringsalgoritme for SRTP er “Advanced Encryption Standard” (AES) i ”counter mode”. Det må etableres en nøkkel som kan brukes til å kryptere RTP-datastrømmen. SRTP spesifiserer ikke prosessen for etablering av nøkler. Se avsnitt 4.5 for mer informasjon om nøkkelhåndtering og nøkkelderivasjon. Etter at nøklene er etablert vil applikasjonen kryptere RTP-nyttelast med AES i ”counter mode”, en 128-bit krypteringsnøkkel og en salt-nøkkel. SRTP-pakkene sendes så over nettet. Bruk av AES i ”counter mode” i SRTP gjør det mulig å prosessere RTP-pakkene selv om de ikke blir mottatt i riktig rekkefølge. Tap av mellomliggende pakker vil ikke ødelegge muligheten for dekryptering av påfølgende pakke.

Standard meldings-autentiseringsalgoritme for SRTP er HMAC-SHA-1 med bruk av en 160-bit nøkkel. ”Hash-based Message Authentication Code” (HMAC) SHA-1 genereres ved å beregne en hash av hele RTP-meldingen, inkludert RTP-header og kryptert nyttelast (mediastrømmen).

Autentisering og integritet av SRTP-pakken er ivaretatt av HMAC-SHA1 nøkkelsignaturen for hver pakke. HMAC-SHA1-verdien blir plassert i en ”Authentication tag” header i SRTP-pakken. Autentisering og integritetsbeskyttelse av RTP-meldinger er viktig for å beskytte mot angrep, som for eksempel ”message replay” og modifikasjon av header og mediastrøm uten at det blir oppdaget [5;6].

RTP-header blir integritetsbeskyttet men ikke kryptert, fordi noen nettverkselementer trenger tilgang til RTP-header (for eksempel ved avregning), i tillegg til at det må være mulig å komprimere RTP-header. Det vil si at RTP-header blir sendt i klartekst og at det bare er RTP-nyttelast som blir kryptert. Angriper kan derfor gjøre trafikkanalyse ved å samle informasjon fra SSRC RTP-headeren, sammen med informasjon fra underliggende transportprotokoller (IP, UDP). SSRC identifiserer avsender (utstyrsenhet) til en RTP-mediastrøm [6].

RTCP blir som tidligere nevnt brukt til å gi QoS-informasjon til deltakere i en multimediasesjon. Sensitiv informasjon som trenger beskyttelse i en RTCP-melding inkluderer informasjon om avsender og innhold av kontrollmeldingene. Disse headerne blir derfor kryptert i SRTCP. Hvis RTCP ikke blir beskyttet kan en angriper manipulere RTCP-meldingene mellom deltakerne av en multimediasesjon eller gjennomføre en trafikkanalyse [6]. Feltet SSRC (utstyrsenhet) blir ikke kryptert i SRTCP og kan dermed fortsatt bidra til trafikkanalyse.

4.3.2 Oppsummering sikkerhet RTP

Det er anbefalt som et minimum å bruke autentisering og integritetsbeskyttelse av RTP-meldinger. Det optimale er å bruke både autentisering, integritetsbeskyttelse og kryptering av meldinger. Standardprotokollen for beskyttelse av sanntids multimediakommunikasjon som tale og video er SRTP.

Fordeler ved SRTP:

- Konfidensialitetsbeskyttelse av nyttelast (mediainnhold) i RTP-pakke og av flere headere og nyttelast i RTCP-pakke.
- Autentisering og integritetsbeskyttelse av hele RTP og RTCP-pakke, både header og nyttelast.
- Gir beskyttelse mot replay-angrep for både RTP- og RTCP-pakker.

Begrensinger ved SRTP:

- Ingen kryptering av RTP-header gir mulighet for trafikkanalyse ved å samle informasjon fra RTP-header
- Kan ikke bevare konfidensialitets-, integritets- og autentiseringbeskyttelse av mediastrøm når den er sendt mellom et IP-nettverk og et PSTN-nettverk
- Nøkkelhåndtering og nøkkelfornying er krevende og påvirker prosessering og ressursforbruk i store multicast-grupper. Dette er ikke ønskelig for mobile enheter som har begrensede beregningsressurser

4.4 Sikring av VoIP med andre sikkerhetsmekanismer

Sikring av VoIP er imidlertid mer enn bare å implementere sikkerhetsmekanismer i SIP og RTP. Siden sikkerhetsdesignet i SIP og RTP er såpass fraværende, anbefales det at man benytter seg av andre sikkerhetsmekanismer for å realisere ønskede sikkerhetsegenskaper. Dette kan være oppsett av IPSec-tunneler mellom VoIP-leverandører, kryptert "Virtual Private Network" (VPN) fra bruker til VoIP-leverandørs gateway, brannvegger og så videre. Her har vi også til dels de samme utfordringene med fordeling av nøkler og kompleksitet som begrenser skalerbarheten.

4.5 Mekanismer for nøkkelhåndtering

Nøkkelhåndtering er en grunnleggende del av beskyttelsen av VoIP. For å kunne beholde konfidensialitets- og integritetsbeskyttelse av mediastrømmen er det viktig med en robust og sikker mekanisme for utveksling av kryptografiske nøkler. Det er to metoder for å utveksle nøkkelmeldinger i SIP- og RTP-baserte VoIP-systemer. Ved bruk av SIP-protokollen (SDP-nyttelast) eller gjennom egne nøkkelutvekslingsmekanismer som MIKEY, ZRTP og SRTP Security Descriptions. Nøkkelutvekslingsmekanismene vil utveksle en hovednøkkel og en Salt-verdi (tilfeldig verdi) som vil bli brukt hos bruker til å utlede sesjonsnøkler for beskyttelse av mediastrømmen.

4.5.1 MIKEY, ZRTP og SRTP Security Descriptions

MIKEY og ZRTP er nøkkelforvaltningsprotokoller som kan brukes for å støtte nøkkelutvekslingen i SRTP. SRTP Security Descriptions er en mekanisme for å forhandle kryptografiske nøkler mellom brukere i unicast-sesjoner som bruker SRTP. Ingen av disse nøkkelutvekslingsmekanismene er valgt til å være industristandard.

MIKEY ("Multimedia Internet KEYing") protokollen er definert i IETF RFC 3830 og ble utviklet for å støtte forhandling av nøkler for sikkerhetsprotokoller som SRTP og IPsec. Standarden beskriver mekanismer for forhandling av nøkler mellom to eller flere parter, som ønsker å etablere en sikker kommunikasjonskanal. MIKEY-meldinger kan bli utvekslet gjennom den signalleringsprotokollen multimedia-applikasjonen bruker. I dette tilfellet vil MIKEY-meldinger bli utvekslet i SDP-nyttelasten i SIP-INVITE og SIP-OK meldingene. SIP må kunne tilby konfidensialitets- og integritetsbeskyttelse av nøklene. Derfor bruker flest implementeringer TLS (TCP), i tillegg til noe utprøving av DTLS (UDP). Eventuelt kan S/MIME brukes for å beskytte SDP-nyttelast.

SRTP Security Descriptions kommuniserer også nøkkelinformasjonen i SDP-nyttelast i SIP-INVITE meldingen. Den er som MIKEY avhengig av at SDP-nyttelast blir kryptert (TLS eller IPsec). Hvis SDP-nyttelast som kommuniserer nøkkelinformasjon ikke blir kryptert vil en angriper kunne snappe opp nøkkelinformasjonen og dekryptere mediastrømmen.

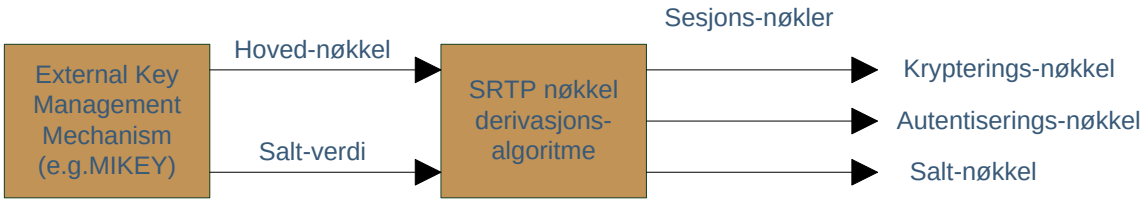
ZRTP-protokollen muliggjør nøkkelforhandling for SRTP ende-til-ende gjennom RTP-mediastrømmen. Ved behov kan ZRTP også forhandle nøkkelmateriell gjennom signalleringsmeldinger. ZRTP-protokollen bruker ephemeral DH (Diffie-Hellman) nøkler til å

etablere en felles hemmelighet mellom endepunktene. ZRTP benytter seg ikke av PKI (det vil si en betrodd 3.part) men baserer seg på et tillitsnettverk. Det vil si at ved bruk av ZRTP kreves det at den enkelte bruker har kontroll over hvilke nøkler den stoler på. En går ikke via en 3.part, men bygger seg opp et tillitsnettverk, også kalt et ”web-of-trust”. Ved å basere seg på ”web-of-trust” omgår ZRTP problemene med en PKI, men skalerer ikke så godt siden den er avhengig av å stole på nøkkelen til den enkelte bruker. ZRTP er implementert i blant annet produktet Zfone, som krypterer VoIP-samtaler ende-til-ende over Internet [21]. Dessverre har ikke ZRTP fått noe industriell momentum, og er og vil trolig forbli et nisjeprodukt for de spesielt interesserte. ZRTP er definert i IETF RFC 6189.

Alle nøkkelutvekslingsprotokoller er begrenset ved at de ikke støtter sesjoner som går mellom VoIP-nettverk og PSTN. For mer informasjon om nøkkelforhandling og nøkkelderivasjon se [6].

4.5.2 Utledning av sesjonsnøkler i SRTP

Nøkkelutvekslingsmekanismene utveksler en hovednøkkel og en Salt verdi (tilfeldig tall) som blir brukt i en nøkkelderivasjonsfunksjon (KDF) hos endebruker for å beregne sesjonsnøkler som skal beskytte RTP-mediastrømmen se Figur 4.2.



Figur 4.2 Nøkkelderivasjon i SRTP [6]

Hver deltager i en RTP-sesjon vil inneha et sett av kryptografisk informasjon for hver SRTP-mediastrøm, som kalles en kryptografisk kontekst. For hver kryptografiske kontekst er det som minimum en krypterings-, autentiserings- og saltnøkkel for henholdsvis SRTP og SRTCP. Salt-nøkkelen blir brukt sammen med krypterings-nøkkelen og krypteringsalgoritmen i SRTP for kryptering av SRTP-meldingene. Autentiseringsnøkkelen blir brukt sammen med autentiseringsalgoritmen for å fremskaffe autentisering- og integritetsbeskyttelse av SRTP-meldingene. I en VoIP-samtale mellom to endepunkter vil det være to SRTP-mediastrømmer, en SRTP-mediastrøm hver vei [6].

5 Oppsummering og konklusjon

VoIP benytter seg av pakkesvitsjet IP-nettverk for transport av signalleringsinformasjon og multimedia. Det er mange forskjellige løsninger for realisering av VoIP-telefoni i dag. De største åpne standardene for realisering av VoIP er SIP og RTP. Oppsett, håndtering og vedlikehold, samt terminering av sesjoner blir håndtert av signalleringsprotokollen SIP. Transport av selve talestrømmen, eller annen multimedia, blir håndtert av protokollen RTP. Det eksisterer i tillegg en god del proprietære løsninger for realisering av VoIP som blir mye brukt, blant annet Skype og

Google Talk. Skype og Google Talk har gode sikkerhetsløsninger med god kryptering. Da de proprietære løsningene er lukket for innsyn fra utenforstående har denne rapporten tatt for seg VoIP basert på SIP og RTP.

Sikkerhet i VoIP- og multimediasesjoner omhandler hvordan VoIP- og multimediatrafikk kan sikres med hensyn til sikkerhetsegenskapene konfidensialitet, integritet og tilgjengelighet. Det er en utbredt oppfatning at VoIP basert på SIP og RTP er svært sårbar for angrep og har et mangelfullt sikkerhetsdesign. Gode sikkerhetsmekanismer kan forhindre flertallet av disse sårbarhetene som er diskutert. Disse har imidlertid foreløpig begrenset utbredelse hos VoIP-leverandører grunnet mangelfull støtte i VoIP-produktene, at de gir økt kompleksitet ved bruk og at de dermed medfører økte kostnader.

Det som brukes av sikkerhet for SIP-signallering i dag er hovedsakelig DAA-autentisering. DAA blir bare benyttet innad i et lokalt domene (typisk en bedrift), og når VoIP-trafikken rutes ut av det lokale domenet benyttes det som oftest ingen sikkerhetsmekanismer i SIP. Utnyttelse av svakheter i DAA kan føre til ”hijacking” av registrering og ”brute-force” angrep for knekking av passord, som kan føre til misbruk av identitet. Andre sikkerhetstrusler grunnet dårlig sikkerhetsdesign i SIP er for eksempel MitM-angrep, misbruk av utgående samtaler, ulovlig avlytting, trafikkanalyse, terminering av sesjoner og DoS-angrep.

Sikkerhetsmekanismer som kan benyttes for å beskytte seg mot truslene nevnt over, er blant annet Secure SIP (SIPS) og S/MIME. Secure SIP gir tosidig autentisering, integritets- og konfidensialitetsbeskyttelse av SIP-meldinger på hopp-per-hopp basis. S/MIME gir ende-til-ende beskyttelse for SIP-meldingen. I praksis blir S/MIME lite brukt. SIPS har vunnet noe støtte hos VoIP-leverandører (industrien) og det ser ut som industrien er i ferd med å velge SIPS som felles sikkerhetsmekanisme. Industrien har sannsynligvis valgt å gå for SIPS fordi TLS er en utbredt industristandard. SIPS kan bli mer utbredt når flere SIP-implementasjoner implementerer dette, samt hvis det kommer en løsning for PKI i telefoni som tilbyr skalerbarhet i forhold til bruk i mange forskjellige telefoner og kommunikasjonsenheter.

Det er anbefalt som et minimum å bruke autentisering og integritetsbeskyttelse av RTP-meldinger. Det optimale er å bruke både autentisering, integritetsbeskyttelse og kryptering av RTP-meldinger. Standardprotokollen for beskyttelse av sanntids multimediakommunikasjon som tale og video er SRTP. Hvis ikke SRTP benyttes, og ingen annen sikkerhetsmekanismer blir benyttet for å beskytte RTP-data, er RTP svært sårbar for både avlytting og manipulasjon. SRTP har foreløpig begrenset utbredelse blant annet på grunn av økt teknisk kompleksitet som SRTP medfører, og at utstyrsprodusentene har vært sene med å implementere SRTP i sine produkter.

Sikring av SIP og RTP basert VoIP vil imidlertid være mer enn kun å implementere sikkerhetsmekanismer i SIP og RTP. Dette kan være oppsett av IPSec-tunneler mellom VoIP-leverandører, ”Virtual Private Network” (VPN) mot endebrukere, brannvegger og så videre. Siden sikkerhetsdesignet i SIP og RTP så langt er såpass fraværende, anbefales det at slike sikkerhetsmekanismer anvendes.

VoIP basert på SIP og RTP er i dag svært sårbar for angrep da gode sikkerhetsmekanismer som SIPS og SRTP i liten utstrekning benyttes. Flere VoIP-leverandører benytter seg av andre sikkerhetsmekanismer for å møte de sikkerhetskrav de måtte ha, men disse har også begrensninger i forhold til kompleksitet og skalerbarhet. Det gjenstår å se om SIPS og SRTP vil få større utbredelse i fremtiden.

Forkortelser

AES – Advanced Encryption Standard
ATA – Analog Telephone Adapter
ATM – Asynchronous Transfer Mode
CA – Certificate Authority
CPU – Central Processing Unit
DAA – Digest Access Authentication
DDoS – Distributed DoS
DH – Diffie-Hellman
DiffServ – Differentiated Services
DNS – Domain Name System
DoS – Denial of Service
DTLS – Datagram Transport Layer Security
ENUM – E.164 NUmber Mapping
FIFO – First In First Out
HD sound – High Definition sound
IETF – Internet Engineering Task Force
IM – Instant Messaging
IMS – IP Multimedia Subsystem
IntServ – Integrated Services
IP – Internet Protocol
IPsec – IP security
IP-telefoni – Internet Protocol-telefoni
ISDN – Integrated Services Digital Network
ISP – Internet Service Provider
LAN – Local Area Network
MAC – Message Authentication Code
MIKEY – Multimedia Internet KEYing
MitM – Man-in-the-middel
MPLS – Multiprotocol label Switching
NAT – Network Adress Translation
PAKE – Password Authenticated Key Exchange
PKI – Public Key Infrastructure
PSTN – Public Switched Telephone Network
PTT – Push-to-talk
RFC – Request For Changes
RTCP – Real-time Transport Control Protocol
RTP – Real-time Transport Protocol
S/MIME – Secure/Multipurpose Internet Mail Extensions
SASL – Simple Authentication and Security Layer
SDP – Session Description Protocol

SIP – Session Initiation Protocol
SIPit 28 – International SIP Interoperability Test Conference 28
SIPS – Secure SIP
SPIT – Spam over Internet Telephony
SRTCP – Secure Real-time Transport Control Protocol
SRTP – Secure Real Time Protocol
TCP – Transmission Control Protocol
TLS – Transport Layer Security
UA – User Agent
UC – Unified Communication
UDP – User Datagram Protocol
URI – Uniform Resource Identifier
VoIP – Voice over IP
VOIPSA – The VoIP Security Alliance
VoLTE – Voice over LTE
VPN – Virtual Private Network
WLAN – Wireless LAN

Referanser

[1] "Det norske ekommarkedet første halvår 2011," Post og Teletilsynet, 2011.

[2] "Skype," http://en.wikipedia.org/wiki/Skype: Wikipedia, 2012.

[3] S. Ganguly and S. Bhatnagar, "VoIP: Wireless, P2P and New Enterprise Voice over IP," Wiley, 2008.

[4] N. Kitawaki and K. Itoh, ""Pure delay effects on speech quality in telecommunications", Selected Areas in Communications," IEEE Journal on, vol.9, no.4, pp.586-593: 1991.

[5] J. Epstien, "Scalable VoIP Mobility:Integration and Deployment," Elsevier, 2009.

[6] P. Thermos and A. Takanen, "Securing VoIP Networks; Threats, Vulnerabilities and Countermeasures," Addison-Wesley, 2007.

[7] O. Hersent, "IP Telephony: Deploying VoIP Protocols and IMS infrastructure," Wiley, 2010.

[8] "Session Initiation Protocol (SIP)," http://en.wikipedia.org/wiki/Session_Initiation_Protocol: 2011.

[9] "SIP:Session Initiation Protocol," IETF RFC 3261: 2002.

[10] L. Strand and W. Leister, "A Survey of SIP Peering," 2011.

[11] "SDP:Session Description Protocol," IETF RFC 4566: 2006.

[12] "RTP:Real-time Transport Protocol," IETF RFC 3550: 2003.

[13] L. Strand, "The development of security architectures in fixed and mobile telephone systems," PhD trial lecture, Ifi, UiO, 2011.

[14] "VoIP Security and Privacy Threat Taxonomy," http://voipsa.org/Activities/taxonomy.php: Voice over IP Security Alliance (VOIPSA), 2005.

[15] L. Strand, "Advancement towards secure authentication in the Session Initiation Protocol," Doctoral Dissertation: Department of Informatics, Faculty of Mathematics and Natural Sciences, University of Oslo, 2011.

[16] L. Strand and A. M. Hagalisletto, "Designing Attacks on SIP Call Setup," International Journal of Applied Cryptography, vol 2, 2010.

[17] E. Sundby Boysen and L. Strand, "Security analysis of the SIP Handover Extension," 2009.

[18] L. Strand and W. Leister, "Advancement towards secure authentication in the Session Initiation Protocol," International Journal on Advances in Security: 2011.

[19] A. P. Hveem, "Mobilt bredbånd med LTE - teknologi, sikkerhet, tjenester og utbygging," FFI-rapport 2011/00709: 11 A.D..

[20] "The Secure Real-time Transport Protocol (SRTP)," IETF RFC 3711: 2004.

[21] P. Zimmermann, www.zfoneproject.com: Zfone Project, 2011.