NTNU
Det skapende universitet

Sikkerhet i VoIP-portnere

David Kristensveen

Master i kommunikasjonsteknologi
Oppgaven levert: Juni 2006
Hovedveileder: Svein Johan Knapskog, ITEM
Medveileder(e): Olav Kvittem, UNINETT

Norges teknisk-naturvitenskapelige universitet
Institutt for telematikk

Oppgavetekst

Det har utviklet seg et marked med et stort antall
IP-telefoni-leverandører med lukkede løsninger som har samtrafikk med hverandre over linje-basert telefoni.
Samtidig utvikles det åpne løsninger basert på SIP og ENUM som også vil kunne bli utbredt parallellt med IP-telefoni-leverandørende.En slik løsning er SIP.EDU for internasjonalt akademisk sektor.

Oppgaven skal undersøke tilgjengelige standarder for samtrafikk mellom IP-telefoni-leverandører, og analysere disse med hensyn på sikkerhet og foreslå løsninger for samtrafikk som gir høy fleksibilitet. Det skal
legges vekt på åpne løsninger.

Oppgaven gitt: 23. januar 2006
Hovedveileder: Svein Johan Knapskog, ITEM

Forord

Dette prosjektet har vært utført våren 2006 ved Norges teknisk-naturvitenskaplige universitet(NTNU), Institutt for telematikk.

Jeg vil takke Olav Kvittem ved UNINETT for hjelp til oppgavebeskrivelse og for gode innspill underveis. En takk går også til faglærer Svein Johan Knapskog i forbindelse med igangsettelse av oppgaven og nyttig veiledning.

Trondheim, juni 2006
David Kristensveen

Innholdsfortegnelse

FORORD ..... **II**

INNHOLDSFORTEGNELSE ..... **IV**

FIGURLISTE ..... **VI**

TABELLISTE ..... **VII**

FORKORTELSER ..... **VIII**

SAMMENDRAG ..... **X**

1 INNLEDNING **1**

1.1 BAKGRUNN 1
1.2 PROBLEMSTILLING 1
1.3 AVGRENSNINGER 2
1.4 OPPSUMMERING 3

2 GRUNNLEGGENDE LØSNINGER FOR SAMTRAFIKK I IP-TELEFONI **4**

2.1 SESSION INITATION PROTOCOL 5
2.2 REAL TIME PROTOCOL 13
2.3 PAKKEBASERT MULTIMEDIAKOMMUNIKASJON(H.323) 15
2.4 PUBLIC SWITCHED TELEPHONE NETWORK 16
2.5 GATEWAY FUNKSJONALITET 24
2.6 PROBLEMSTILLINGER I FORBINDELSE MED ADRESSERING 30
2.7 TELEFONNUMRE, IP-ADRESSER OG RUTING 37
2.8 SESSION BORDER CONTROLLERS 51

3 ULIKE MODELLER FOR SAMTRAFIKK MELLOM BEDRIFTER **58**

4 PEER-TO-PEER LØSNINGER FOR IP-TELEFONI OVER INTERNETT **62**

4.1 SKYPE 62
4.2 SIP SOM PEER-TO-PEER 63

5 THE IP MULTIMEDIA SUBSYSTEM OG 3GPP **64**

5.1 ARKITEKTUR 64
5.2 SAMTRAFIKK 66
5.3 SIKKERHET 69
5.4 NYE TJENESTER I IMS 71

6 UNLICENSED MOBILE ACCESS **72**

6.1 ARKITEKTUR 72
6.2 SIKKERHET 74

7 UTFORDRINGER OG KRAV TIL SIKKERHET **77**

7.1 SIKKERHET I FORBINDELSE MED SIP 77
7.2 SPAM OVER IP-TELEFONI 79
7.3 SIKKERHETSMEKANISMER I FORBINDELSE MED TRÅDLØSE NETTVERK 82
7.4 TJENESTEKVALITET 88
7.5 SIKKERHET I FORBINDELSE MED ENUM OG DNS 90

8 DISKUSJON **94**

8.1 SAMTRAFIKK 95
8.2 ENUM BASERT PÅ E.164.ARPA (PUBLIC ENUM) 96
8.3 INFRASTRUKTUR ENUM 97

8.4 IP-TELEFONI OVER TRÅDLØSE DATANETTVERK 98
8.5 VIDERE UTVIKLING 99
8.6 KONKLUSJON 101
9 REFERANSER 102

Figurliste

Figur 2.1 ”Eksempel på en SIP melding” modifisert fra [12] 9
Figur 2.4 ”Real Time Protocol i kombinasjon med UDP og IP” modifisert fra [36] 13
Figur 2.4.1 “Nettverkstopologien for signalering i SS7” modifisert fra [40] 17
Figur 2.4.2 “Lagdelt oversikt over funksjonaliteten i SS7” modifisert fra [40] 18
Figur 2.4.3 “Meldingsflyt med metoden *All Call Query* for ruting” modifisert fra [37] 20
Figur 2.4.4 “Meldingsflyt med metoden *Query on Release*” modifisert fra [37] 21
Figur 2.4.5 “Meldingsflyt med metoden *Call Dropback*” modifisert fra [37] 22
Figur 2.4.6 “Meldingsflyt med metoden *Onward routing*” modifisert fra [37] 23
Figur 2.5 ”Samtrafikk mellom IP-telefoni og PSTN” 24
Figur 2.5.1 “Oppbygging av en enkel Media Gateway” 25
Figur 2.5.2 ”Oversikt over utviklingen av ulike Gatewayprotokoller” fra [9] 26
Figur 2.5.2.2 “Samarbeid mellom MGC og MG” modifisert fra [41] 27
Figur 2.5.3 “Logisk oppbygging av en gateway som benytter MGCP” modifisert fra [7] 28
Figur 2.5.3.1 ”Konferanse mellom tre parter ved bruk av Megaco” modifisert fra [7] 28
Figur 2.6.1 ”Logisk oppsett av signalering og mediatransport” modifisert fra [1] 30
Figur 2.6.2 ”Klient bak NAT kommuniserer med en STUN-server” modfisert fra [1] 34
Figur 2.6.3 ”Klient bak NAT kommuniserer med en TURN-server” modifsert fra [1] 35
Figur 2.7.1 ”Overordnet arkitektur av domener som benytter TRIP” modifisert fra [45] 38
Figur 2.7.2 ”Telefonisamtale fra PTSN til SIP ved hjelp av ENUM” 42
Figur 2.7.3 ”Lagdelt struktur i forbindelse med ENUM” modifisert fra [24] 43
Figur 2.7.4 “Registreringsprosses i DNS i forbindelse med ENUM” modifisert fra [24] 44
Figur 2.7.5 “Standard oppsett av samtale i en SIP.edu arkitektur” modifisert fra [38] 46
Figur 2.7.6 “Utvidet oppsett av samtale i en SIP.edu arkitektur” modifisert fra [38] 47
Figur 2.8.1 ”Plassering av enkel SBC mellom to nettverk” modifisert fra [46] 51
Figur 2.8.2 ”Plassering av delt SBC mellom to nettverk” modifisert fra [46] 52
Figur 2.8.3 “SBC mellom tjenestetilbyder og kunde” modifisert fra [46] 53
Figur 2.8.4 ”Eksempel på Via-feltet i SIP” fra [1] 54
Figur 2.8.5 ”REGISTER melding med SBC” modifisert fra [46] 57
Figur 3.2.1 ”Atskilte nettverk for tale og data” modifisert fra [21] 58
Figur 3.2.2 ”IP-telefoni for interne samtaler og PSTN for eksterne” modifisert fra [21] 59
Figur 3.3.3 ”Løsninger med bruk av gateway” modifisert fra [21] 60
Figur 3.3.4 ”Løsninger med bruk av ENUM” modifisert fra [21] 61
Figur 5.1 ”Sammenheng mellom ulike komponenter i IMS” 64
Figur 5.2. ”Samtrafikk mellom to IMS enheter” modifisert fra[6] 66
Figur 5.2.1 “Samtale fra IMS over til linjesvitsjet nettverk” modifisert fra[18] 67
Figur 5.2.2 “Gateway funksjonalitet i IMS” modifisert fra[6] 68
Figur 5.2.3 “Samtale fra linjesvitsjet nettverk og inn i IMS” modifisert fra[6] 68
Figur 5.3 ”Autentisering i IMS” modifisert fra [9] 69
Figur 5.3.1 ”Generering av autentiseringsvektor i IMS” fra [47] 70
Figur 6.1 ”Mobiltelefoni ved hjelp av UMA” modifisert fra[16] 72
Figur 6.1.1 ”Enheter som kommuniserer med UMA-nettverkskontroller” fra[17] 73
Figur 6.2 ” Sikkerhetsforbindelser i forbindelse med mobiltelefoni” modifisert fra [17] 74
Figur 7.3 ”Bruk av delt verdi for autentisering” 82
Figur 7.3.1 ”Klientautentisering ved hjelp av EAP” 84
Figur 7.3.2 ”EAP-AKA i forbindelse med UMA” fra [17] 85
Figur 7.4 “Bruk av Differentiated Services i forbindelse med SIP” 89
Figur 7.5 ”Sikkerhetsmekansimer i forbindelse med ENUM” modifisert fra[25] 90

Tabelliste

Tabell 2.4.2 ”Oversikt over ISUP meldinger” 19
Tabell 2.6.1 ”Portnumre for mediatstrømmen ved hjelp av innholdet i SDP” 31
Tabell 6.2 “Ulike sikkerhetsprofiler i forbindelse med UMA” modifisert fra[17] 76
Tabell 7.1.1 ”Ulike sikkerhetsmekanismer i forbindelse med SIP 77
Tabell 7.1.2 ”Ulike typer angrep som kan rettes mot SIP” 78
Tabell 7.3.1 ”Grad av sikkerhet i forbindelse med ulike sikkerhetsmekanismer” 82
Tabell 7.3.2 ”Varianter av Extended Authentication Protocol(EAP)” modifisert fra [9] 86
Tabell 8.1 ”Sluttbruker ENUM og Infrastruktur ENUM” modifisert fra[23] 97

Forkortelser

- ALG (Application Level Gateway): Programvare som benyttes sammen med brannmurer for å analysere trafikken på applikasjonsnivået.
- E.164: Navnet på den internasjonale telefonnummerplanen administrert av ITU (International Telecommunication Union). Et fullt kvalifisert E.164-nummer inneholder en landkode, en by- eller områdekode pluss et telefonnummer. I Norge inngår by eller områdekoden i selve telefonnummeret.
- ENUM(E.164/Telephone Number Mapping): Oversettelse av et E.164 telefonnummer til et domenenavn.
- IMS(The IP Multimedia Subsystem): Rammeverk som skal sørge for en ytterligere sammensmelting av data- og linjesvitsjede nettverk. Signaleringen innad i IMS er basert på SIP.
- MG(Media Gateway): Skal konvertere strømmen av media mellom IP-nettet og PSTN–nettet.
- MGC(Media Gateway Controller): Har ansvaret for å samkjøre Media Gateway og Signaling Gateway slik at både samtaleoppsettet og samtalen kan gjennomføres.
- MGCP(Media Gateway Control Protocol): Benyttes for å kommunisere mellom de ulike partene i en dekomponert gateway.
- NAT(Network Address Translation): Metode for å forandre på kilde- eller mottakeradressen til en IP-pakke.
- RTP (Real Time Protocol): Standard som benyttes til transport av sanntidstrafikk.
- RTCP(Real Time Control Protocol): Benyttes for å analysere RTP-trafikken.
- SDP (Session Description Protocol): Format for å utveksle parametere i forbindelse med oppsett av multimediasesjoner.
- SG(Signaling Gateway): Benyttes å realisere signalering mellom PSTN-nettet og IP-nettet.
- SIP(Session Initiation Protocol): Benyttes til å initiere, modifisere og terminene interaktive sesjoner.

- SS7 (Common Channel Signaling System #7): Arkitektur som benyttes i forbindelse med signalaring i forbindelse med oppsett av sesjoner, betalingstjenester, ruting og andre funskjoner i forbindelse med PSTN.
- TRIP(Telephony Routing over IP): Protokol som benyttes for å utveksle telefon-rutinginformasjon mellom administrative domener.
- UMA(Unlicensed Mobile Access): Teknologi som benyttes for å tilby kunder mobiltelefoni over 802.11 eller Blutetooth.
- WEP(Wired Equivalent Privacy) Den originale standarden som ble beskrevet i forbindelse med autentisering og kryptering i forbindelse med 802.11.
- WPA(Wi-Fi Protected Access): Viderutvikling av WEP

Sammendrag

Session Initiation Protocol(SIP) er i ferd med å bli den ledende signaleringsprotokollen i forbindelse med IP-telefoni. SIP benyttes til å initiere, modifisere og terminere interaktive sesjoner. Arkitekturens to hovedkomponenter er servere og brukeragenter. De ulike brukeragentene utveksler forespørsler og tilhørende responsmeldinger. Etter at en sesjon er satt opp av SIP benyttes Real Time Protocol(RTP) til å overføre data i forbindelse med selve samtalen. RTP benytter dynamisk valgte portnumre. Disse utveksles på forhånd mellom brukeragentene ved hjelp av Session Description Protocol(SDP) i meldingskroppen til SIP-meldingene.

Når brannmurer eller Network Address Translation(NAT) benyttes sammen med IP-telefoni er det et som regel et problem at IP-adresser og portnumrene som skal benyttes av brukeragentene omskrives. Adresseinformasjonen som er utvekslet på forhånd vil derfor bli ugyldig. Det finnes flere ulike metoder for å tilnærme seg disse problemene. To metoder fra IETF er Simpel Traversal of UDP through NAT(STUN) og Traversal using Relay NAT(TURN). Session Border Controllers(SBC) er lukkede kommersielle nettverksløsninger som benyttes til å løse mange av de samme problemene. Ulempen med SBC er at dette er kostbare løsninger og at prinsippet om at SIP skal være en åpen protokoll brytes.

Innefor IP-telefoni ser en utvikling der tilbydere eller organisasjoner har infrastrukturen på plass for å realisere en IP-telefonitjeneste på vegne av sine egne brukere, mens samtrafikk med andre er nødt til å foregå ved hjelp av PSTN-nettet. For å realisere en slik samtrafikk er en nødt til å benytte en Gateway. En slik Gateway har oftest en todelt funksjonalitet. Først må det oversettes mellom signaleringen i IP-telefoni(SIP) og signaleringen i PSTN(som regel ISUP). Dette utføres av en signaliseringsgsgateway(SG). Deretter må mediastrømmen oversettes av en mediagateway(MG) fra RTP til det aktuelle formatet som benyttes av det linjesvitsjede nettverket. En Media Gateway Controller(MGC) benyttes for å samkjøre MG og SG. Telephony Routing over IP(TRIP) er en protokoll som kan benyttes av tjenestetilbydere eller organisasjoner for å utveksle rutingtabeller for sine respektive gatewayer. En annen mekanisme for PSTN til IP

samtrafikk er SIP for Telephones(SIP-T). Her kan PSTN-signaleringen enten oversettes til SIP eller pakkes inn i SIP-meldinger.

En av hovedgrunnene til at PSTN benyttes til ruting i IP-telefon skyldes at det ikke har eksistert noen fullgod erstatter for Signaleringssystem nummer syv(SS7) i forbindelse med IP-telefoni. E.164 Number Mapping(ENUM) er en metode for å oversette E.164-numre til domenenavn ved hjelp av DNS. Ved en slik løsning vil en være i stand til å realisere samtrafikk mellom ulike typer IP-telefoninettverk uten å benytte PSTN. I motsetning til tradisjonelle løsninger der mange i dag benytter to ulike nettverk for data- og taletjenester vil en fremover kunne se en utvikling der alle tjenester er basert på IP og PSTN vil bli overflødig. Samtidig ser en utvikling der Internett løsninger som Skype blir stadig mer populære. Det er også muligheter for å benytte SIP til tilsvarende løsninger, men her er det fortsatt en del uenighet i om hvordan slike løsninger konkret skal realiseres.

The IP Multimedia Subsystem(IMS) er et rammeverk som skal sørge for ytterligere konvergens mellom telefoni- og datatjenester. Telefonikunder skal tilbys nye multimediatjenester på apllikasjonsnivå. Kjernenettverket i IMS skal benytte SIP-baserte løsninger. IMS omtales ofte som neste generasjons nettverk.

Etter hvert som 802.11 nettverk er blitt mer utbredt er det et voksende marked for IP-telefoni over 802.11. Unlicensed Mobile Access(UMA) er en teknologi som lar brukeren benytte mobiltelefonen over 802.11 eller Bluetooth. Samtidig tillates det handover mot det mobile nettverket når dette er nødvendig. Sikkerhet i trådløse nettverk er et område under stadig utvikling. Ved innføring av 802.11i vil en få sterkere mekanismer for sikkerhet i form av Extended Authentication Protocol(EAP) for autentisering og Advanced Encryption Standard(AES) for kryptering.

Ved introduksjonen av ENUM oppstår det nye sikkerhetsutfordringer. Et system med billigere telefoni og enkel tilgangs til brukerlokasjoner vil potensielt kunne føre til mer Spam over IP-telefoni(SPIT). Ellers er DNS systemet utsatt for mange potensielle trusler.

1 Innledning

I dette kapitelet vil det bli gitt en beskrivelse av bakgrunnen for oppgaven. Deretter vil det bli beskrevet hvilke problemstilinger det har vært fokusert på, og hvilke avgrensninger som er gjort. Til slutt gis en kort gjennomgang av innholdet i de ulike kapitlene.

1.1 Bakgrunn

Samtrafikk mellom ulike tilbydere av IP-telefoni er i dag ofte nødt til å rutes over PSTN-nettet. For ruting i PSTN-nettet har signalering i forbindelse med SS7 vært mest benyttet, og det har ikke eksistert noen fullgod erstatter for SS7 i forbindelse med IP-telefoni. Tilbyderne ha derfor basert seg på egne lukkede løsninger internt for å holde oversikten over lokasjonsinformasjon for sine brukere. På grunn av manglende mekanismer for å dele denne informasjonen har en vært nødt til å benytte PSTN for ekstern ruting.
Slike løsninger medfører ekstra kostnader for de ulike tilbyderne. Dette skyldes i hovedsak betalingskostnader til de ulike PSTN-tjenestetilbyderne. I tillegg er det også lite lønnsomt å benytte to ulike typer nettverk(IP og PSTN) for en tjeneste som i utgangspunktet skal fungere kun ved hjelp av IP. Dette medfører at endebrukerne må betale en høyere pris enn nødvendig for telefonitjenensten.

1.2 Problemstilling

PSTN-nettverket kontrolleres i dag av tradisjonelle telefonitilbyderne. Ved siden av dette har det etter hvert kommet flere tilbydere, organisasjoner eller bedrifter med sine egne høyhastighets IP-nettverk. Ved en teknologi som ENUM vil disse være i stand til å utveksle rutinginformasjon for sine brukere, og en vil derfor kunne unngå ruting via PSTN-nettet. Oppgaven vil belyse ulike problemstillinger i forbindelse med denne formen for samtrafikk.

Ved siden av dette vokser det fram flere kommersielle sammensmeltinger av IP-telefoni, applikasjonstjenester og tradisjonelle telefonitjenester. To løsninger for dette er IP Multimedia Subsytem(IMS) og Unlicensed Mobile Access(UMA). Oppgaven gir en overordnet oversikt over disse to teknologiene. I tillegg vil det bli drøftet noen sikkerhetsaspekter i forbindelse med disse ulike temaene.

1.3 Avgrensninger

I oppgaven presenters det ulike løsninger som kan benyttes for å muliggjøre samtrafikk i IP-telefoni. Det gis også en del eksempler på utfordringer som oppstår i forbindelse med disse løsningene. Oppgaven inneholder ikke noen ”kokebokoppskrift” for hvordan en skal realisere en komplett IP-telefonitjeneste, men fokuset har isteden vært å belyse sentrale problemstillinger. Når det gjelder signalering vil det bli i hovedsak bli satt fokus på SIP, da det synes å være en bred enighet om at SIP i framtiden vil være den dominerende signaleringsprotokollen i forbindelse med IP-telefoni.

Teknologiene IMS og UMA er løsninger som er drevet av kommersielle aktører. IMS er allikevel interessant på den måten at SIP er valgt som signaleringsprotokoll, og at mange av spesifikasjonene er basert på åpne løsninger fra IETF. Kapitelet om UMA er tatt med for å belyse en trend der mobiltelefoner også kan støtte bruk av IP-telefoni.

Når det gjelder språket er det stor sett benyttet engelske faguttrykk der dette faller naturlig, dette gjelder spesielt uttrykk som inngår i mye brukte forkortelser. Blant annet er ordet *Gateway* benyttet i stedet for det norske ordet *Portner*. I flertall benyttes gatewayer istedenfor gateways.

I mange eksempler i denne oppgaven skilles det mellom IP-telefoni og PSTN i forbindelse med samtrafikk og bruk av ENUM. Her er PSTN da brukt som et samlebegrep for både PSTN, ISDN og GSM basert telefoni.

Høsten 2005 skrev jeg prosjektoppgaven ”Sikkerhetsutfordringer ved IP-telefoni”[10]. Mange av temaene der er også aktuelle i denne masteroppgaven. For å unngå å gjenta meg selv for mye har jeg i denne oppgaven prøvd å fokusere på emner som jeg ikke allerede har skrevet om i den forrige oppgaven. Dette medfører at det i denne oppgaven ikke fokuseres så mye på kun sikkerhet, men på mer på generelle problemstillinger i forbindelse med IP-telefoni.

1.4 Oppsummering

Kapitel 2: Dette er hovedkapitlet og grunnlaget for resten av oppaven. Her gjennomgåes det en del grunnleggende løsninger for samtrafikk i IP-telefoni. Dette innebærer blant annet en gjennomgang av ulike meldinger i forbindelse med SIP og PSTN samt problemstillinger i forbindelse med ruting og adressering.
Kapitel 3: Her blir det gitt en gjennomgang av hvordan ulike bedrifter eller organisasjoner kan bygge opp sine ulike nettverk i forbindelse med IP-telefoni.
Kapitel 4: Inneholder en kort beskrivelse av Peer-to-Peer løsninger over Internett.
Kapitel 5: Dette kapitelet omhandler The IP Multimedia Subsystem(IMS). Det blir gitt en overordnet beskrivelse av arkitekturen. Det blir også gjennomgått et par eksempler i forbindelse med samtrafikk mot PSTN.
Kapitel 6: Dette kapitelet gir en beskrivelse av Unlicensed Mobile Access(UMA). Det er fokusert på overordnet arkitektur og sikkerhet.
Kapitel 7: Her gjennomgåes en del problemstillinger i forbindelse med sikkerhet i IP-telefoni og samtrafikk.
Kapitel 8: Dette kapitelet er en diskusjon i forhold til emner som tideligere er omtalt i oppgaven.
Kapitel 9: Inneholder oversikt over referanser.

2 Grunnleggende løsninger for samtrafikk i IP-telefoni

I dette kapitlet vil det bli gjennomgått en del ulike områder i forbindelse med samtrafikk i IP-telefoni. Først vil det bli gitt en gjennomgang av Session Initiation Protocol(SIP) og Session Description Protocol(SDP). Grunnen at dette er vektlagt skyldes at SIP er i ferd med å bli den dominerende signaliseringsprototokollen i forbindelse med IP-telefoni, og mange av eksemplene i resten av kapitlet er basert på arkitekturer som benytter SIP.

Det vil også bli gitt en kort oppsummering av signalering og arkitektur i forbindelse med Public Switched Telephone Network(PSTN). For å realisere samtrafikk mellom SIP og PSTN kreves det en arkitektur som benytter ulike typer Gatewayer, en beskrivelse av dette vil derfor bli gjennomgått.

Problemstillinger i forbindelse med Network Address Translation(NAT) er sentrale i forbindelse med IP-telefoni. Dette er spesielt knyttet opp mot problemstillinger i forbindelse med ruting. Det vil derfor bli presentert en del ulike tilnærminger for hvordan en kan kombinere IP-telefoni med NAT.

Stadig flere tilbydere og organisasjoner ønsker å sørge for samtrafikk seg imellom uten å benytte PSTN-nettet. Med nye teknologier som ENUM er en blitt mer uavhengig av PSTN-nettet. ENUM og andre løsninger for ruting vil derfor bli gjennomgått. Tilslutt vil det bli presentert en løsning som kalles Session Border Controllers(SBC). SBC er kommersielle nettverksenheter som benyttes for å løse mange av problemstillingene med samtrafikk, dette gjelder spesielt håndtering av NAT.

2.1 Session Initation Protocol

Først i dette kapitelet vil det bli blitt gitt et kort sammendrag av viktige elementer i arkitekturen som benyttes av Session Initation Protocol(SIP). Noe av dette stoffet har jeg tideligere omtalt i forbindelse med prosjektoppgaven min høsten 2005[10]. Siden dette er viktig grunnstoff også for denne oppgaven vil kapittel 2.1.2 inneholde en kort oppsummering vedrørende brukeragenter, ulike typer servere og meldingsflyt i forbindelse med SIP sesjoner. I denne oppgaven vil det for øvrig bli fokusert mer på innholdet i meldingene og ikke bare den overordnete arkitekturen.

2.1.1 Funksjonalitet

Kort fortalt så benyttes SIP til å initiere, modifisere og terminene interaktive sesjoner. I tillegg til dette ble det under utviklingen av SIP stilt fem konkrete krav som protokollen skulle innfri:

1) Brukerlokasjon. Det skal være enkelt å finne ut hvor endebrukeren er lokalisert. Dette innebærer at brukernavnet skal kunne oversettes til en IP-adresse.
2) Brukertilgjengelighet: Brukere av SIP skal ha muligheten til selv å bestemme om de ønsker å bli oppfattet som tilgjengelige eller ikke.
3) Brukerstyrte parametere: Brukerne skal selv kunne sette opp parametere i forbindelse med mediaoverføringen. Et eksempel på dette kan være at brukeren bestemmer om det er en IP-telefonisamtale eller en videooverføring som skal initieres.
4) Initiere sesjoner: Brukerne skal selv ha muligheten til å sette opp sesjoner.
5) Administrere sesjoner: Brukerne skal selv ha muligheten til å styre sesjonen. I praksis betyr dette tjenester som overføring av samtaler, sette en samtale på vent og avslutning av samtaler.

2.1.2 Overordnet arkitektur

Arkitekturen som benyttes i forbindelse med SIP kan deles inn i to hovedkomponenter:

1) SIP Brukeragenter: Dette er endepunktene i samtalen. Brukeragentene handler på vegne av brukerne og beskrives ved hjelp av klient/server arkitektur. Brukeragenten som sender en forespørsel er klient og brukeragenten som blir forespurt er server.
2) SIP Servere: En brukeragent må registrere seg hos en server. Serveren holder så oversikt over tilstandsinformasjon, brukernavn og IP-adresse for brukeragenten. Det finnes ulike servere med ulike funksjoner. Vi skiller mellom proxyservere, redirectservere og registerservere.

Adresser

En SIP adresse er bygd opp på samme måte som en e-post adresse med brukerID og et vertsnavn. BrukerID kan enten være et brukernavn eller en E.164 adresse. Vertsnavnet kan være et domenenavn eller en nettverksadresse. Betegnelsen URI (Uniform resource identificator) benyttes for å beskrive disse adressene.

Registerserver

Denne serveren holder oversikt over lokasjonen til brukeragentene som er pålogget på nettverket. I praksis er dette som oftest en database som holder rede på bindingen mellom brukernavnet og den tilhørende IP-adressen.

Proxyserver

En proxyserver brukes til å videresende forespørsler på vegne av brukeragentene. Proxyserveren kan også benyttes til aksesskontroll, autentisering og autorisasjon.

Redirectserver

Redirectserveren tilbyr adresseinformasjon til brukeragenten slik at brukeragenten selv kan sende forespørsler. Forskjellen på en redirect- og en proxyserver er altså at proxyserveren videresender forespørsler på vegne av klienten, mens redirecetserveren overlater dette ansvaret til brukeragenten.

2.1.3 Ulike typer meldinger

Meldingene som sendes av en brukeragent kan enten være en forespørsel eller en respons på en forespørsel. I RFC3261[11] deler derfor IETF opp SIP-meldingene i de to kategoriene forespørsel og respons. Som sagt så omtales brukeragenten som sender forespørsler for klient, og brukeragenten som responderer kalles for server. Denne rollefordelingen mellom brukeragentene vil fortløpende veksle under en sesjon avhengig av hvem som sender hvilke typer meldinger.

SIP forespørsel

IETF definerer i [11] seks ulike typer forespørsler. Disse er.

1) INVITE: Benyttes for å initiere en sesjon mellom to brukeragenter.
2) REGISTER: Brukeragenten sender denne forespørselen til en Registerserver når brukeragenten skal oppdatere serveren om sin lokasjon.
3) ACK: Etter at brukeragenten har initiert en sesjon med en annen brukeragent, og fått respons på denne forespørselen så sendes en ACK melding for å bekrefte at sesjonen kan starte.
4) BYE: Benyttes for å avslutte en sesjon. Denne meldingen skal kun sendes etter initiativ fra brukeragentene. I en åpen SIP-arkitektur skal ikke mellomliggende servere på egenhånd ha muligheten for å avslutte sesjoner mellom to brukeragenter.
5) CANCEL: Brukes for å avbryte et sesjsonsoppsett etter at en INVITE melding er sendt fra brukeragent A. CANCEL kan kun brukes i påvente av respons fra brukeragent B. Hvis brukeragent B allerede har respondert på INVITE forespørselen må BYE benyttes istedenfor CANCEL.
6) OPTIONS: Benyttes på samme måte som INVITE, men formålet er ikke å initiere en sesjon men å undersøke om en brukeragent er tilgjengelig.

SIP respons

Disse meldingene blir benyttet av brukeragenten som fungerer som server. SIP benytter seks ulike typer responsmeldinger. SIP bygger på mange av de samme klient/server prinsippene som Hypertext Transfer Protocol(HTTP) og de responsmeldingene som

benyttes av SIP deles inn på samme måte som i HTTP. SIP benytter seks klasser for responsmeldinger. Disse listes opp i eksemplet under, sammen med noen konkrete meldinger som er relevante for IP-telefoni og som benyttes i del eksempler utover i oppgaven.

1) 1XX(Informerende): Brukes for oppdatering av status underveis i oppsettet av en sesjon.
    - 180 ”Ringing”: Varsler brukeragent A om at INVITE-forespørselen er mottatt av brukeragent B, og at vedkommende varsles om dette.
    - 182 ”Samtalen er satt i kø”: Varsler brukeragent A om at INVITE-forespørselen er mottatt, men forespørselen er satt i kø.
2) 2XX(Suksess): Varsler brukeragent A om at forespørselen er akseptert av brukeragent B.
    - 200 ”OK”: Sendes fra brukeragent B til brukeragent A for å varsle om at forespørselen aksepteres. Samtidig oversendes ulike parametere for mediaoverføringen. Dette er nærmere beskrevet i kapittel 2.1.4.
3) 3XX(Forflytning): Disse meldingene sendes som regel ut fra Redirectserveren til brukeragent A som sendte INVITE forespørselen for å oppdatere om brukeragent B sin adresse.
    - 301 ”Moved permanently”: Varsler brukeragent A om brukeragent B sin nye permanente adresse.
    - 302 ”Moved temporarily”: Varsler brukeragent A om brukeragent B sin midlertidige adresse.
4) 4XX(Feil i klient): Disse meldingene mottas av brukeragent A som prøve å initiere en sesjon. De kan sendes ut fra en server eller brukeragent B for å varsle om feil.
    - 400 ”Bad request”: Varsler om at serveren ikke klarte å tolke forespørselen.
    - 404 ”Not found”: Serveren klarer ikke å lokalisere brukeragenten som forespørres.
5) 5XX(Feil i server): Meldingen sendes fra serveren til brukeragenten som prøver å initiere en sesjon. Meldingen sier at forespørselen ikke kan prosesseres på grunn av feil i serveren.

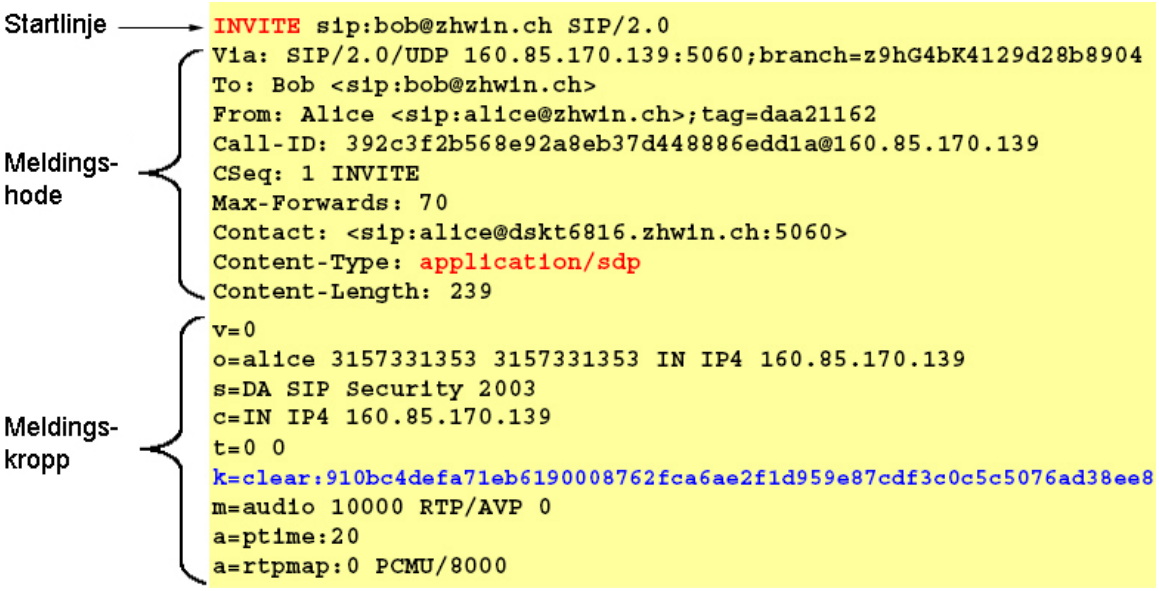
6) 6XX(Global feil): Disse meldingene finnes ikke i HTTP og er konstruert spesielt med tanke på SIP.

- 600 ”Busy everywhere”: Varsler brukeragent A og at brukeragent B er for tiden er utilgjengelig og ikke kan nåes fra noen lokasjoner.

2.1.4 Innholdet i en melding

SIP er en tekstbasert protokoll som er basert på ASCII koding. En SIP-melding består av en starlinje, meldingshode(header) og en meldingskropp(body). Oppbygging av meldingshodet i en SIP melding er den samme som benyttes i Simple Mail Transfer Protocol(SMTP). I forbindelse med innholdet i meldingskroppen benyttes Session Description Protocol(SDP).

Figur 2.1 ”Eksempel på en SIP melding” modifisert fra [12]



Startlinje/Statuslinje

Startlinjen inneholder informasjon vedrørende den aktuelle forespørselen/responsen. I en forespørsel har startlinjen formen(metode, mottaker-URI, versjon). I en respons kalles startlinjen ofte for statuslinje og har formen(versjon, statuskode) og i tillegg et valgfritt tredje felt som kan brukes til å beskrive responsen.

Meldingshode

Meldingshodet inneholder informasjon relatert til selve meldingen. Her beskrives det hvem som skal kontaktes, hvem som sender forespørselen og en egen ID for samtalen. Følgende felter er obligatoriske:

- VIA: Brukes for å oppdatere adresseinformasjon ettersom meldingen transporteres over nettet. Dette er nærmere omtalt i kapittel 2.8.
- To: Indikerer hvem som er mottageren av meldingen.
- From: Indikerer hvem som har dannet meldingen.
- Call-ID: Unik identifikasjon for hver enkelt samtale.
- Cseq: Sekvensnummer som økes med verdien en for hver melding. På denne måten kan de to brukeragentene holde oversikt om meldingene kommer i riktig rekkefølge.
- Max-Forwards: Indikerer maksimalt antall steg mellom avsender og mottaker.
- Contact: Inneholder adressen som brukeragenten ønsker å motta meldinger på.

Meldingskropp og Session Description Protocol

For at en IP-telefonisamtale skal kunne gjennomføres er brukerne i begge ender avhengig av å benytte en del felles parametere. Dette gjelder for eksempel talekoding, valg av mediaprotokoll og hvilke portnummer som skal benyttes til samtalen. Session Description Protocol(SDP) benyttes i meldingskroppen til en SIP-melding for å overføre disse parametrene vedrørende den aktuelle SIP-sesjonen som skal settes opp.

I figur 2.1 ser vi et eksempel på en SIP INVITE melding som benytter SDP i meldingskroppen. For at innholdet skal kunne analyseres enklest mulig representeres ulike type informasjon med en liten bokstav. Følgende felter er definert som obligatoriske:

- Version(”v”): Beskriver hvilken versjon av SDP som benyttes.
- Owner(”o”): Beskriver hvem som eier sesjonen. Dette er brukeren som initierer sesjonen. Feltet skal være unikt for hver enkelt bruker. Fra figur 2.1 kan vi se:
    - Brukernavn: alice
    - Identifikasjon: 3157331353

    - Versjonsnummer: 3157331353
    - Meldingen skal sendes over Internet(IN) med IP versjon 4 og brukers IP-adresse er 160.85.170.139.

  Identifikasjonsnummeret og versjonsnummer dannes ved hjelp av et tidsstempel. Identifikasjonsnummeret er konstant under hele sesjonen, mens versjonsnummeret oppdateres ved endringer i sesjonen.
- Session name(”s”): Inneholder navnet på sesjonen.
- Connection information(”c”): Inneholder nettverkstypen som skal benyttes, hvilken adressetype som skal benyttes og den konkrete adressen til avsender.
- Time(”t”): Inneholder to verdier. Disse beskriver start- og sluttid på sesjonen. Ved å benytte verdien null for sluttiden indikeres det at det ikke er satt noen begrensninger på varigheten. Dette er den mest vanlig løsningen fordi det i de fleste tilfeller ikke er naturlig å sette tidebegrensninger på en sesjon.

I tillegg til de obligatoriske feltene er det også en mange felter som kan velges utfra hvilke behov som gjelder for den aktuelle sesjonen. Noen av de meste benyttede valgfrie feltene er:

- Key(”k”): Her oversendes den krypteringsnøkkelen som avsender ønsker at skal benyttes for å kryptere den kommende sesjonen. Dette emnet er nærmere omtalt i kapittel 3.6 i prosjektoppgaven min ”Sikkerhetsutfordringer ved IP-telefoni”[10].
- Media(”m”): Fortelle hva slags media som skal benyttes i sesjonen og hvordan dette skal overføres. Fra figur 2.1 ser vi at det valgte mediet er lyd og at overføringen skal foregå på port 10000 ved hjelp av Real Time Protocol. Den valgte meidaprofilen for overføringen er 0.
- Attributt(”a”): Dette er attributtfelter som kan benyttes for å gi et ytterligere beskrivelse av sesjonen. Disse feltene er nærmest ment som hjelpefelter og kan droppes i analysen av SDP-innholdet hvis de ikke kan tolkes av mottaker. Fra figur 2.1 ser vi at opplysningen beskriver at pakkelengden er 20 ms og at mediaprofilen for overføringen er 0.

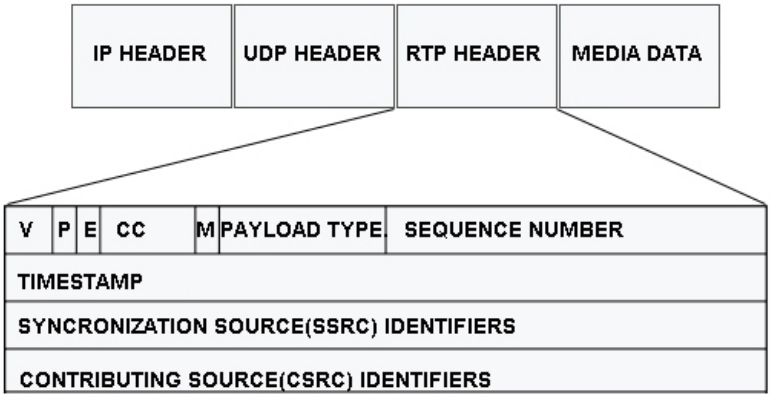
2.1.5 SIP Back-to-Back User Agents

En Back-to-Back User Agent(B2BUA) er en SIP-enhet som mottar en forespørsel fra en brukeragent og omskriver denne før den sender forespørselen videre. Responsen på denne forespørselen omskrives også før den sendes tilbake til brukeragenten. En av nytteverdiene ved en slik tjenestene sett fra en tilbyders synspunkt er at brukeragenten sin adresseinformasjon skjules ovenfor den andre brukeragenten. Dette realiseres ved at B2BUA lager nye felter for områdene Via, From, Contact, CallID og omskriver innholdet i meldingskroppen. De modifiserte feltene er nå omskrevet slik at de peker tilbake på B2BUA og ikke brukeragenten. Å implementere en ordning med B2BUA har den ulempen at det bryter med prinsippet om at SIP skal være en åpen protokoll. Noe av ansvaret for samtaleoppsettet flyttes over fra brukeragentene og over til sentraliserte nettverksenheter. I verste fall kan en B2BUA bli en flaskehals i et system siden denne er nødt til å oversette all trafikken mellom brukeragentene. Oppetiden for en tjeneste som benytter B2BUA kan også bli redusert siden tilgjengeligheten til en rekke brukeragenter er avhengig av denne ene enheten. B2BUA er også implementert i mange moderne brannmurer, men da omtales de ofte som Application Layer Gateways(ALG).

2.2 Real Time Protocol

Etter at en sesjon er satt opp mellom to brukeragenter ved hjelp av signaleringsprotokoller som SIP eller H.323 kan mediautvekslingen(talen) starte. Som kjent så overføres talen i en IP-telefoni samtale ved hjelp av Real Time Protocol(RTP). RTP er designet fra grunnen med hensyn på å overføre mediastrømmer i sanntid. RTP benyttes sammen med User Datagram Protocol(UDP). RTP-pakkene fraktes i nyttedelen av en UDP-pakke.

Figur 2.4 "Real Time Protocol i kombinasjon med UDP og IP" modifisert fra [36]



På toppen av figur 2.4 er den en logisk fremstilling av hvordan RTP-pakkene kombineres med UDP og IP. Nederst i figuren er innholdet i RTP-headeren uthevet. Disse feltene er nødvendige for å illustrere virkemåten for RTP. Det er særlig fire felt som er viktige:

1) Sekvensnummer: Dette nummeret økes inkrementelt med verdien en hver gang en ny RTP-pakke er sendt. Det er derfor mulig for mottaker å registrere om pakkene kommer i riktig rekkefølge og om pakker er blitt borte underveis.
2) Tidsstempel: Hver RTP-pakke gis et tidsstempel. Akseptabel forsinkelse for en RTP-pakke i forbindelse med IP-telefoni er 150ms. I tillegg kan det oppstå variasjoner i forsinkelsen mellom hver enkelt pakke. Dette kalles "jitter" og medfører at en samtale vil oppleves som hakkete. Ved hjelp av tidsstempelet kan en korrigere "jitter" på mottakersiden
3) Synkroniseringskilde identifikasjon: I en forbindelse som er satt opp med RTP kan det komme RTP-pakker fra ulike kilder, som har ulike krav til synkronisering.

Hvis det i en forbindelse benyttes både telefoni og kamera vil disse ha forskjellig synkroniseringsidentifikasjon.

4) Bidragskilde identifikasjon: Ved en konferanse hvor det sitter fem personer i rom A med hver sin mikrofon og prater med en person i rom B kan det være lønnsomt å mikse lyden fra de 5 ulike mikrofonene inn i samme kanal. For at personen i rom B skal kunne skille de 5 ulike mikrofonen benyttes en bidragskilde identifikasjon.

2.2.1 Real Time Control Protocol

Real Time Control Protocol(RTCP) benyttes for å gi avsender eller mottaker tilbakemeldinger vedrørende trafikken som transporteres med RTP. RTCP pakkene sendes periodevis til deltakerne i en RTP-sesjon og benytter samme distribusjonsform som RTP-pakkene. Alle deltakere i en sesjon sender RTCP pakker og informasjonen kan blant annet benyttes til flytkontroll og holde oversikt over om deltakere har forlatt en RTP-sesjon.

2.2.2 Bruk av portnumre

Portnummeret som benyttes i forbindelse med RTP bør ifølge spesifikasjonen alltid være et partall. I en del operativsystemer er de ikke mulig å velge et fast portnummer for en applikasjon som benytter RTP. Dette skyldes kollisjoner mellom ulike RTP-applikasjoner som kjører på den samme maskinen. Portnummeret velges derfor som regel tilfeldig av applikasjonen som skal benytte RTP.

Portnummeret som skal benyttes til RTCP-pakkene skal alltid være et nummer høyere enn det valgte portnummeret for RTP. Siden portnummeret for RTP alltid skal være et partall vil portnummeret for RTCP alltid bli et oddetall.

Problemstillinger i forbindelse med portnumre er nærmere omtalt i kapitel 2.6.

2.3 Pakkebasert multimediakommunikasjon(H.323)

I [10] har jeg omtalt grunnleggende arkitektur og en del sikkerhetsmekanismer i forbindelse med H.323. Siden hovedfokuset i denne oppgaven er på SIP ville ikke dette temaet vil ikke bli omtalt i denne oppgaven. For lesere som ønsker å vite mer om dette kan jeg anbefale å lese min forrige oppgave eller ITU sin egen spesifikasjon[39]. I denne oppgaven har jeg heller valgt å kort belyse noen elementære forskjeller på SIP og H.323.

Forskjeller på SIP og H.323

H.323 har sin bakgrunn fra ITU, mens SIP er utviklet av IETF. H.323 ble konstruert for å fungere sammen med PSTN og har derfor arvet flere elementer fra PSTN, blant annet innenfor koding av meldinger og signalering. SIP er i utgangspunktet konstruert med tanke på å fungere over Internett og metodene i SIP er derfor mer rettet mot samtrafikk over Internett.

De to viktigste bruksområdene for H.323 i dag er som en ren erstatning for PSTN-nettet og innenfor videokonferanser. For basis telefonitjenester uten noen behov for tilleggstjenester har nettverk som allerede benytter H.323 lite å tjene på å gå over til SIP.

SIP er som kjent en tekstbasert protokoll som er basert på elementer fra HTTP og SMTP H.323 benytter ASN.1 kodede meldinger. Dette medfører at meldingene blir mer komprimert, men samtidig mer kompliserte og vanskeligere å implementere. SIP sin enkle ASCII baserte koding har vært en av grunnene til at SIP stadig økende popularitet. SIP-meldinger er enkle å analysere og det er lett å skrive programvare for simulering, testing og analyse. H.323 er kun en signaleringsprotokoll. SIP er også en signaleringsprotokoll, men den kan i større grad utvides med ny funksjonalitet. Dette gjelder blant annet tjenester som markering for forskjellige typer av tilstedværelse som vi kjenner fra MSN Messenger. SIP vil også spille en sentral rolle i utviklingen av “The IP Multimedia Subsystem” som er nærmere omtalt i kapittel fem.

2.4 Public Switched Telephone Network

I dag er PSTN fortsatt verdens største kommunikasjonssystem, og innenfor tale kommer PSTN til å være dominerende i enda noen år. I denne oppgaven er PSTN først og fremst interessant i forbindelse med samtrafikk med IP-telefoni. Det vil derfor bli en kort gjennomgang av adressering, signalering og ruting i PSTN.

2.4.1 Den internasjonale nummerplanen E.164

Den internasjonale nummerplanen er basert på en spesifikasjonen E.164 fra ITU-T. Alle land som skal følge internasjonal samtrafikk er avhengige av å følge E.164. I denne nummerplanen kan ingen nummer være lenger enn 15 siffer. De tre først sifrene er satt av til landskoder, men en trenger ikke å benytte alle tre. De resterende 12 sifrene representerer de interne numrene innenfor hvert enkelt land.

2.4.2 Signalering

For å sette opp en telefonsamtale kreves det flere ulike former for signalering som samtaleoppsett, rutinginformasjon og betalingsinformasjon. Flere ulike metoder benyttes innen signalering i forbindelse med PSTN. De to mest vanlige er Common Channel Signaling(CCS) og Common Associated Signaling(CAS).

Common Associated Signaling

Signaleringen i CAS foregår i de samme kanalene som taletrafikken. Dette kalles derfor ”in-band” signalering. Dette foregår ved at enkelte av bit’ene i rammene for taletrafikken er dedikert til signalering. Denne metoden er også kjent som ”bit robbing”.

Common Channel Signaling

CSS benytter en egen kanal dedikert til signalering og derfor regnes CSS som ”out of band” signalering. Selv om CSS benytter en egen kanal multiplekses denne som regel sammen med kanalene som er dedikert til taletrafikk og signaleringen transporteres langs den samme linjen som talesignalene. I USA har CSS også vært kjent gjennom varianten Common Channel Interoffice Signaling(CCIS). CCIS er nesten det samme som ITU-T sitt Common Channel Signaling System #6(SS6). Etter hvert som behovet for et signaleringssystem med støtte for samtrafikk mellom tele- og datatrafikk er kommet,

benyttes nå stort sett Common Channel Signaling System #7 (SS7) isteden for CCIS og SS6.

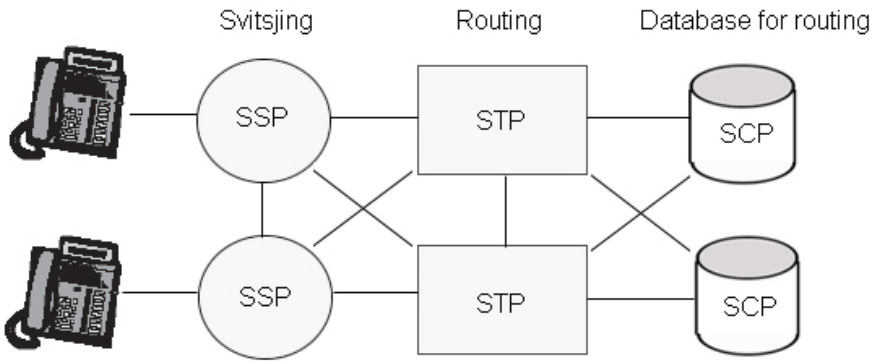
Integrated Services Digital Network

Integrated Services Digital Network(ISDN) ble utviklet for å kunne benytte digitale telefoner i PSTN. Signalering i ISDN er en videreføring av CSS teknologien fordi også ISDN benytter en egen dedikert kanal til signalering. Basic rate interface(BRI) og Primary rate interface(PRI) er de to mest vanlige variantene av ISDN. BRI benytter to 64Kbs B-kanaler til å transportere tale eller data og en 16Kbs D-kanal til signalering. PRI benytter 23 B-kanaler og en D-kanal til signalering.

Common Channel Signaling System #7 (SS7)

Et nettverk som benytter SS7 er bygget opp av tre hovedkomponenter. Dette er Service Switching Points(SSP), Service Transfer Points(STP) og Service Control Points(SCP).

Figur 2.4.1 "Nettverkstopologien for signalering i SS7" modifisert fra [40]



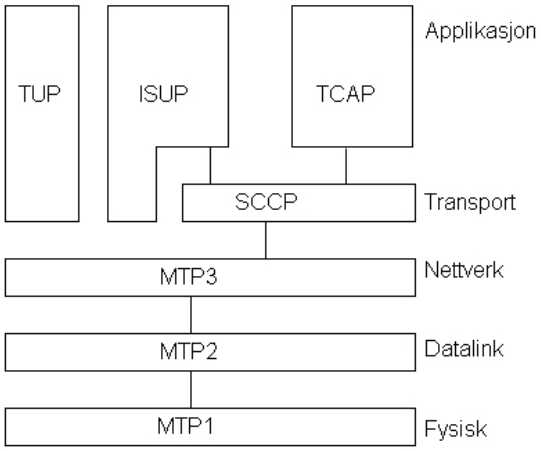
SSP står for svitjsefunksjonaliteten i nettverket. SSP benyttes til basisfunksjoner som oppsett av samtaler, samtalekontroll og avslutning av samtaler. Hver SSP har ansvaret for et gitt område av nettverket. Hvis samtaler skal rutes utenfor dette nettverket er SSP nødt til å kontakte en SCP for å finne ut hvordan samtalen skal rutes.

SCP er en database som svarer på rutingforespørsler fra de ulike SSP'ene. Funksjonaliteten til SCP minner derfor en del om hvordan DNS benyttes i IP-nettverk.

Forskjellen er at det er mye strengere krav til autentisering og autorisasjon for å benytte informasjon som ligger lagret i SCP. Kun autoriserte tjenestetilbydere gis muligheten til å forandre på innholdet i SCP, og hver enkelt SSP må autentiseres for å benytte informasjonen i SCP.

STP ligger sentralt i nettverket og benyttes til ruting innenfor SS7 og kan derfor sammenlignes med en IP-ruter i Internett. På grunn av dette trenger det ikke å være direkte linker mellom de ulike SSP'ene.

Figur 2.4.2 "Lagdelt oversikt over funksjonaliteten i SS7" modifisert fra [40]



Telephone User Part(TUP)

Dette er den eldste varianten. Denne baserer seg på analog signalering. Før signaleringen ble digitalisert ble TUP benyttet til å initiere og avslutte samtaler. I dag har denne varianten stort sett blitt erstattet av ISDN User Part(ISUP).

Transaction Capabilities Applications Part(TCAP)

TCAP støtter signalering for tjenester som ikke er linjesvitsjet, blant annet autentisering og roaming for mobiltelefoni. For å realisere dette er TCAP nødt til å benytte seg av Signaling Connection Control Part(SCCP).

ISDN User Part(ISUP)

ISUP er den delen av SS7 som har ansvaret for ISDN signalering. ISUP definerer de prosedyrene som trengs for å sette opp, administrere og avslutte sesjoner som benyttes for å overføre tale eller datatrafikk over PSTN. ISUP kan også brukes til signalering for samtaler som ikke benytter ISDN, og er i dag den mest brukte formen for signalering i SS7.

Tabell 2.4.2 "Oversikt over ISUP meldinger"

Forkortelse	Melding	Forklaring
IAM	Initial Adresse Message	Benyttes for å initiere samtalen
ACM	Adress Complete Message	Bekreftelse fra mottaker på at samtale kan settes opp
CPG	Call progress	Sendes fra mottaker for å varsle om progresjon i oppsett av samtalen
ANM	Answer message	Bekrefter at mottaker har besvart samtalen
REL	Release	Kan sendes fra begge parter for å avslutte en samtale
RLC	Release complete	Kan sendes fra begge parter for å bekrefte at samtalen er avsluttet.

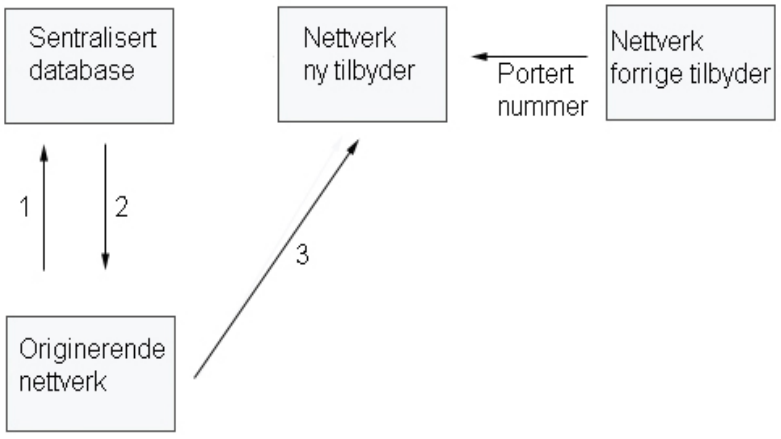
2.4.3 Nummerportabilitet og PSTN

I forbindelse med samtrafikk er det et interessant spørsmål hvordan eierskap av numrene og nummerportabilitet kan gjennomføres. I dette kapittelet vil det derfor presenters noen ulike løsninger fra IETF på hvordan dette kan løses. Det skilles mellom ulike typer nummerportabilitet. Den formen som skal gjennomgås her er såkalt tilbyderportabilitet. Tilbyderportabilitet innebærer at kunden kan beholde sitt telefonnummer, selv om kunden velger å bytte tilbyder av telefonitjenesten. Forutsetningen er at den nye tilbyderen er villig til å tilby denne tjenesten ovenfor sine kunder. På engelsk omtales dette som Service provider number portability(SPNP) og defineres i RFC3482[37]. Bakgrunnen for å innføre nummerportabilitet var at telemarkedene skulle liberaliseres med hensyn på å styrke nye aktører og dermed sørge for en friere konkurranse.

IETF skisserer fire ulike modeller for hvordan samtaler til porterte nummer kan rutes i ett nettverk. Her benyttes betegnelsen donornettverk, dette er en betegnelse på det nettverket som først fikk tildelt det aktuelle telefonnummeret.

All Call Query (ACQ)

Figur 2.4.3 "Meldingsflyt med metoden *All Call Query* for ruting" modifisert fra [37]

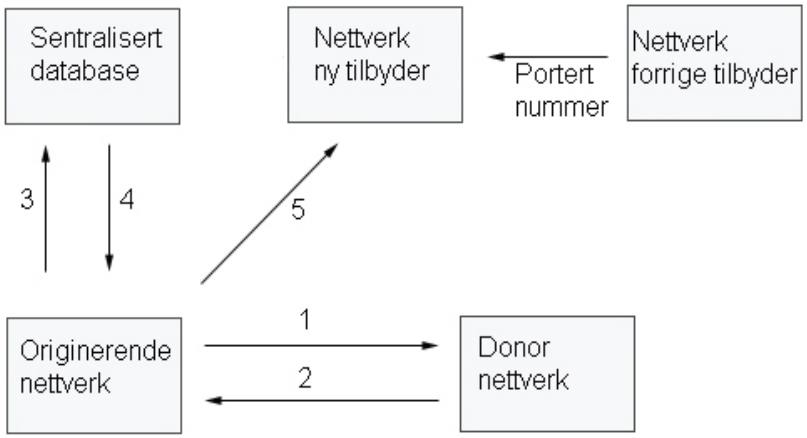


1) Når det originerende nettverket mottar en samtale sendes det en forespørsel til en sentralisert nummerporteringsdatabase (eller en kopi av denne).
2) Databasen svarer med å sende rutinginformasjon om det forespurte nummeret.

3) Det originerende nettverket bruker så denne informasjonen til å sende forespørselen videre til det den oppringte brukeren sitt hjemmenettverk.

Query on Release (QoR)

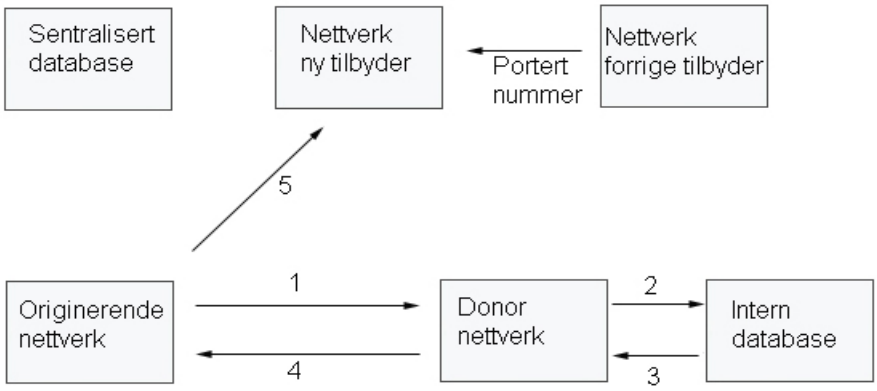
Figur 2.4.4 "Meldingsflyt med metoden *Query on Release*" modifisert fra [37]



1) Når det originerende nettverket mottar en samtale videresendes denne samtalen til donornettverket.
2) Hvis nummeret som forespørres er portert vil donornettverket ikke lenger ha ansvar for rutingen, og forespørselen droppes.
3) Det originerende nettverket må derfor forespørre porteringsdatabasen om rutinginformasjon om det aktuelle nummeret.
4) Databasen svarer med å sende rutinginformasjon om det forespurte nummeret.
5) Det originerende nettverket bruker så denne informasjonen til å sende forespørselen videre til det den oppringte brukeren sitt hjemmenettverk.

Call Dropback

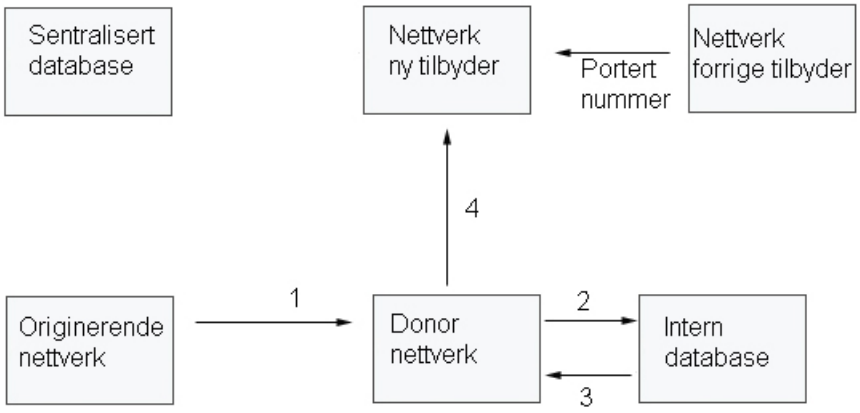
Figur 2.4.5 "Meldingsflyt med metoden *Call Dropback*" modifisert fra [37]



1) Når det originerende nettverket mottar en samtale videresendes denne samtalen til donornettverket.
2) Donornettverket registrerer at det tideligere har vært ansvarlig for dette nummeret. Isteden for å droppe forespørselen som i QoR gjør nå donornettverket ett oppslag i sin interne database for å finne ut hvilket nettverk som nå er ansvarlig for det aktuelle nummeret.
3) Databasen returnerer rutinginformasjon om det aktuelle nummeret.
4) Donornettverket videresender informasjonen til det originerende nettverket.
5) Det originerende nettverket bruker så denne informasjonen til å sende forespørselen videre til det den oppringte brukeren sitt hjemmenettverk.

Onward Routing (OR)

Figur 2.4.6 "Meldingsflyt med metoden *Onward routing*" modifisert fra [37]

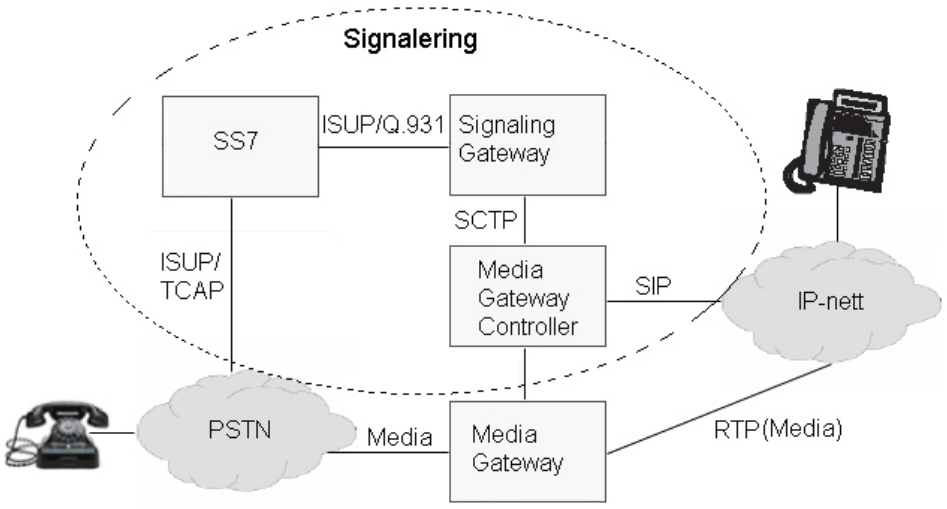


1) Når det originerende nettverket mottar en samtale videresendes denne samtalen til donornettverket.
2) Donornettverket registrerer at det tideligere har vært ansvarlig for dette nummeret. Isteden for å droppe forespørselen som i QoR gjør nå donornettverket ett oppslag i sin interne database for å finne ut hvilket nettverk som nå er ansvarlig for det aktuelle nummeret.
3) Databasen returnerer rutinginformasjon om det aktuelle nummeret.
4) Donornettverket bruker rutinginformasjonen til å rute samtalen til det nye hjemmenettverket til den oppringte brukeren.

2.5 Gateway funksjonalitet

En VoIP gateway fungerer som en adapter mellom datanettet og telenettet og har mange forskjellige funksjoner. Den skal kunne ta imot samtaler og viderekoble samtaler. Samtidig må det kunne gjennomføres en transformasjon mellom analoge og digitale signaler.

Figur 2.5 "Samtrafikk mellom IP-telefoni og PSTN"



Fra utsiden kan en gateway oppfattes som en enkelt enhet, som kan kontaktes via en bestemt adresse. Til tross for dette kan en gateway gjerne være et stort distribuert system. Det er derfor naturlig å dekomponere gateway innholdet inn i tre ulike områder for funksjonalitet.

1. Media Gateway (MG): Skal konvertere strømmen av media mellom IP-nettet og PSTN–nettet. Dette skyldes at talen i PSTN-nettet blir må transporteres som IP-pakker i datanettverket.
2. Signaling Gateway(SG): Er nødvendig for å realisere signalering mellom PSTN-nettet og IP-nettet. Dette kan for eksempel være å transformere ISUP signaler fra PSTN-nettet til SIP på IP-nettet.
3. Media Gateway Controller (MGC): Omtales ofte som samtaleagent. Har ansvaret for å samkjøre MG og SG slik at både samtaleoppsettet og samtalen kan gjennomføres.

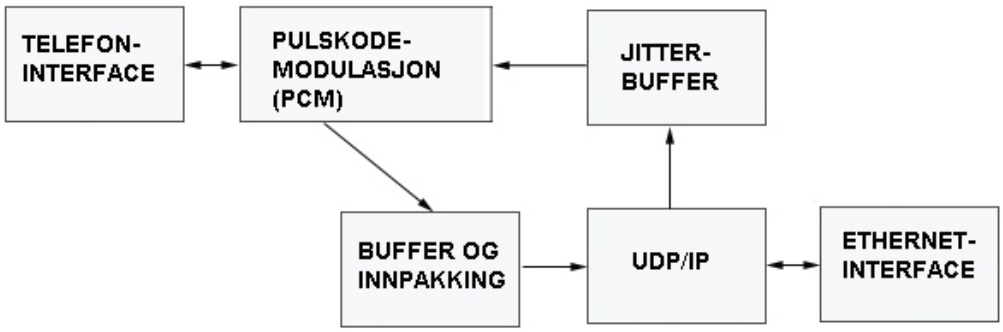
2.5.1 Media Gateway

Hvordan en mediagateway skal fungere kan løses på mange forskjellige måter. I sin enkleste form inneholder en gateway et ethernet-interface i den ende enden og et analogt eller digital telefon-interface i den andre enden.

Analog til digital konvertering

Den enkleste formen for en mediagateway er et analogt-adapter. Figur 2.5.1 viser en logisk oppbygging av et slikt adapter. Omformingen mellom det analoge og det digitale signalet foretas ved hjelp av pulskodemodulasjon

Figur 2.5.1 "Oppbygging av en enkel Media Gateway"



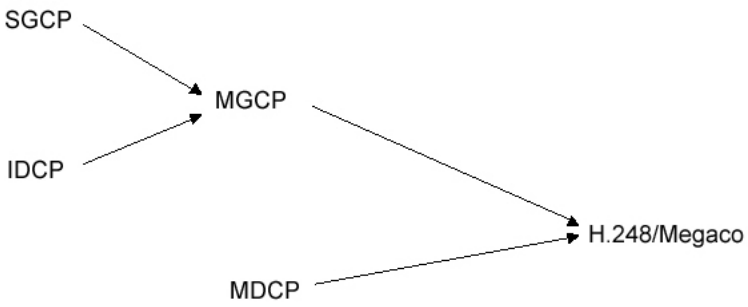
En mye brukt standard for koding/dekoding(KODEK)er G.711.
Her leses det analoge signalet av 8000 ganger per sekund og representeres med 8 bit per avlesning og bitraten blir derfor 64 kbit/s. Siden bitstrømmen skal transporters ved hjelp av IP må den bufres og deles inn i IP-pakker. Dette medfører ekstra overhead og talestrømmen en vei krever derfor omlag 100 kbit/s. Før omgjøring tilbake fra digitale til analog signaler må pakkene først gjennom et jitterbuffer siden de kan ankomme med variert forsinkelse.

2.5.2 Protokoller for koordinering innad i en gateway.

En gateway vil bestå av mange media gatewayer. Hver enkelt media gateway(MG) vil være et kontaktpunkt med omverdenen. Siden uliker MG'er kan være atskilt på ulike adresser kan en gateway betraktes som et distribuert system med svitsjefunksjonalitet. En Media Gateway Controller(MGC) kontrollerer en eller flere slike MG'er. MGC skal sørge for at ulike MG'er kommuniserer seg imellom slik at det kan bli satt opp en forbindelse gjennom nettverket.

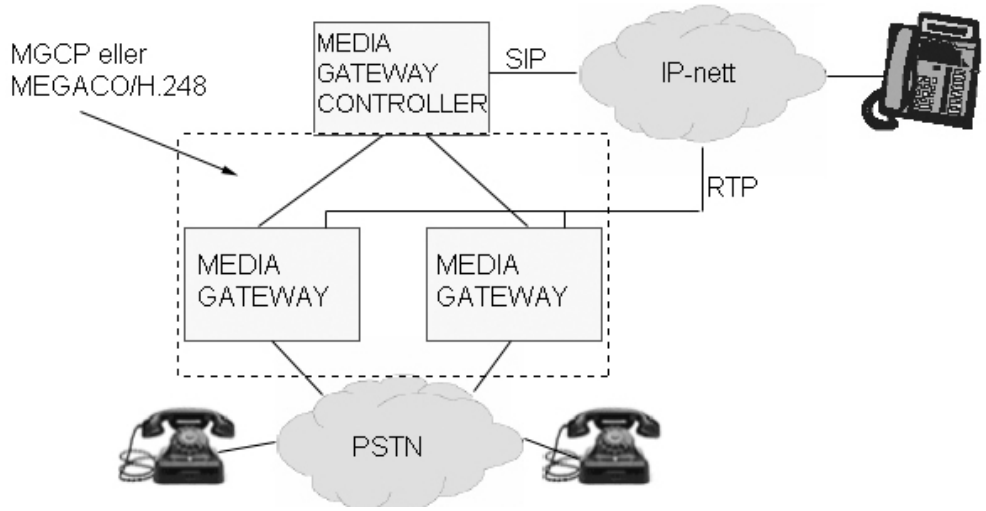
Det er derfor nødvendig med et rammeverk for å standardisere hvordan kommunikasjonen mellom en MGC og en MG skal foregå. Til dette formålet er det utviklet ulike protokoller både fra IETF og ITU.

Figur 2.5.2 "Oversikt over utviklingen av ulike Gatewayprotokoller" fra [9]



I perioden 1998 til 2000 jobbet IETF og ITU med hver sine tilnærminger til Gatewayprotokoller. IETF benyttet Media Gateway Controll Protocol(MGCP) mens ITU benyttet Media Device Control Protocol(MDCP). I 2000 ble det enighet om en felles standard. Denne kalles H.248 av ITU og MEGACO av IETF. Siden navn er lettere å huske enn nummer vil benevnelsen MEGACO bli brukt i denne oppgaven. Siden MGCP fortsatt er mye brukt vil virkemåten til denne bli gjennomgått sammen med virkemåten til MEGACO.

Figur 2.5.2.2 “Samarbeid mellom MGC og MG” modifisert fra [41]



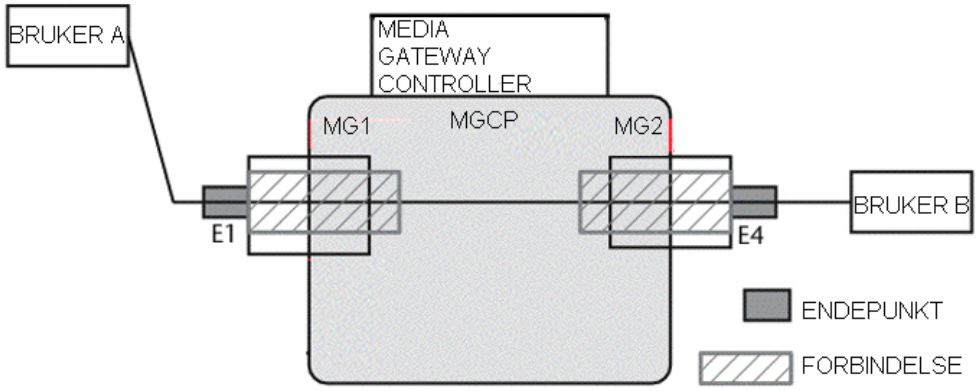
MGCP og MEGACO har den samme overordnede strukturen, men en noe forskjellig funksjonalitet.

Media Gateway Control Protcol

Media Gateway Control Protcol(MGCP) er spesifisert i RFC 3435[42] fra IETF. I denne arkitekturen tar en utgangspunkt i at samtalekontroll er lagt til en egen samtaleagent(media gateway controller). Samtaleagenten benytter SIP eller H.323 for å kommunisere med IP-telefoninettet og SS7/ISUP for å kommunisere med PSTN-nettet.

MGCP benyttes for å kommunisere mellom de ulike partene i en dekomponert gateway. Hver enkelt media gateway holder ikke selv oversikt over gjeldende status for en samtale, men utfører bare kommandoer den får av samtaleagenten. Data utveksles direkte mellom de involverte mediagatewayene. Dette er IP-adresser, portnumre og mediaparametre. Denne informasjonen utveksles ved hjelp av SDP på en tilsvarende måte som i SIP.

Figur 2.5.3 "Logisk oppbygging av en gateway som benytter MGCP" modifisert fra [7]

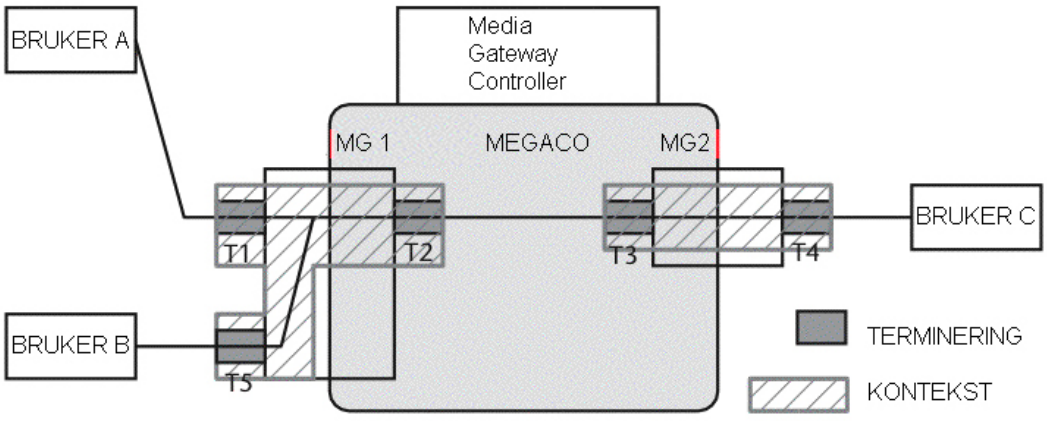


MGCP benytter begrepene endepunkter og forbindelse. Et endepunkt er kontaktpunktet ut mot eller inn fra omverdenen mens forbindelsen beskriver forholdet mellom de ulike endepunktene innad i en gateway.

MEGACO

Megaco er etterfølgeren etter MGCP og inneholder en del mer avanserte muligheter. Siden MEGACO er et samarbeid mellom IETF og ITU støttes både tekstbasert og binær koding. I tillegg støttes muligheter for bruk av TCP for kommunikasjon mellom komponentene innad i en gateway. Samtidig har MEGACO støtte for flere deltakere innefor ulike konferansetjenester.

Figur 2.5.3.1 "Konferanse mellom tre parter ved bruk av Megaco" modifisert fra [7]



I MEGACO benyttes begrepet termineringer istedenfor endepunkter. Det defineres en terminering på begge sider av en media gateway. De ulike termineringene plassers i en kontekst. Fra figuren ser vi at hvis bruker A forlater samtalen vil ikke dette påvirke forbindelsen mellom bruker B og bruker C. Denne oppbyggingen gir MEGACO en mer avansert funksjonalitet enn MGCP. Ved bruk av for eksempel en konferanseserver må en i MGCP sette opp en ende-til-ende forbindelse for hver enkelt deltaker. Ved bruk av MEGACO kan disse forbindelse plasseres i en gitt kontekst som gir flere muligheter. For eksempel kan en ha en konferanse der en gruppe har rett til og prate og lytte, mens en annen gruppe er passive deltakere og kun har rettigheter til å lytte.

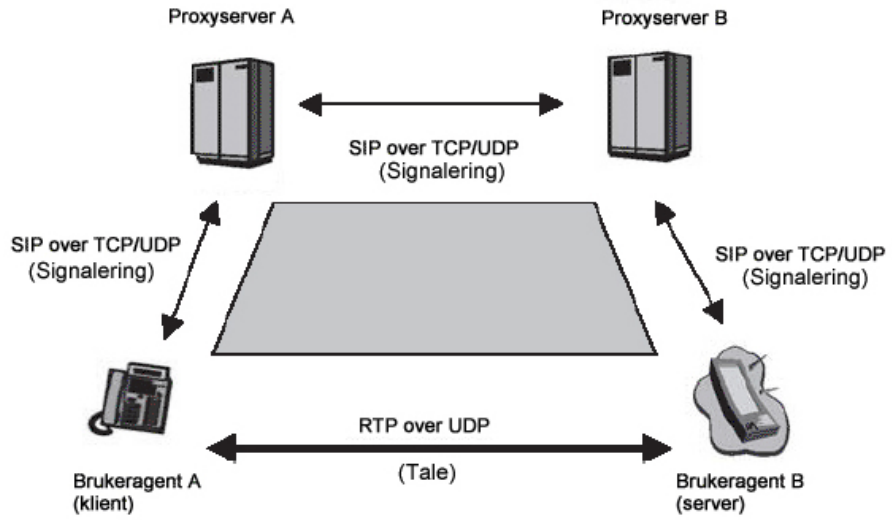
2.6 Problemstillinger i forbindelse med adressering

I dette kapitelet vil det bli gjennomgått noen problemstillinger som oppstår i forbindelse med IP-telefoni når en brukeragent befinner seg i et internt nettverk som benytter Network Address Translation(NAT) for å kommunisere med det eksterne nettverket. Først vil det bli gitt en forklaring på hvordan Network Address Translation fungerer og hvorfor dette er et problem for IP-telefoni. Deretter vil det bli gjennomgått noen ulike tilnærminger for å løse disse problemstillingene.

2.6.1 Bruk av portnumre på transportlaget for SIP og RTP

Som kjent så benyttes SIP til å sette opp sesjoner mellom to brukeragenter, mens RTP benyttes til å overføre selv mediastrømmen.

Figur 2.6.1 "Logisk oppsett av signalering og mediatransport" modifisert fra [1]



I forbindelse med oppsett av SIP-sesjoner så ser vi at før selve mediastrømmen(talen) mellom to brukeragenter kan starte må de første ha en signaleringsfase der de utveksler adresseinformasjon og mediaparametre. Til denne signaleringen kan SIP over Transport Control Protocol(TCP) eller User Datagram Procol(UDP) benyttes, men talen selve talen transporteres med RTP over UDP. Signalering i SIP benytter portnummer 5060 for TCP eller UDP og portnummer 5061 hvis Transport Layer Security(TLS) benyttes. Denne informasjonen sendes i VIA-feltet i headeren på SIP-forespørslene.

TCP og UDP har begge to headerfelt som er satt av til portnumre. Disse to kalles kilde(source) og mål(destination). Disse portnumrene benyttes til å identifisere applikasjonprosesser. Hver portnummer består av 16-bit og kan variere fra 0 og opp til 65535. Portnumrene mellom 0 og 1023 er reservert til kjente applikasjoner som for eksempel http(80) og ftp(21).

Når en klient sender en forespørsel til en server benyttes da mål-portnummeret for å beskrive hvilke applikasjon klienten ønsker å forespørre. Kilde-portnummeret er et tilfeldig valgt portnummer som genereres av klienten. Når serveren så sender en respons tilbake byttes disse to feltene plass. Mål-pornummeret er nå klienten sitt genererte nummer, mens kilde-portnummeret er applikasjonsnummeret på serveren. SIP signalering benytter nå ofte symmetrisk signalering. Det vil si at samme portnummer benyttes både til å sende og motta trafikk.

Som vist så benytter altså SIP velkjente portnumre til signaleringen. RTP benytter derimot dynamisk valgte portnumre for hver enkelt sesjon. Hvilke portnumre som skal benyttes i forbindelse med den aktuelle sesjonen utveksles mellom de to brukeragentene ved hjelp av innholdet i SDP som vist i kapittel 2.1.4. Fra kapitelet om RTP vet vi at kun partall skal benyttes som portnummer i forbindelse med RTP.

Tabell 2.6.1 ”Portnumre for mediatstrømmen ved hjelp av innholdet i SDP”

	SDP parameter	Port kilde	Port mål
Brukeragent A	m=audio 49170 RTP/AVP 0	49170	52310
Brukeragent B	m=audio 52310 RTP/AVP 0	52310	49170

I tabellen ser vi hvordan mediaparametre i SDP benyttes til å utveksle portnumre og informasjon om hvilken protokoll som skal benyttes. Tallet ”0” tilstutt forteller hvilke talekodning som skal brukes. ”0” er en peker til formatet G.711.

2.6.2 Network Address Translation

Network Address Translation(NAT) er en metode for å forandre på kilde- eller mottakeradressen til en IP-pakke når de passerer en NAT-enhet. I dette ligger også den sikkerhetsfordelen at oppbygging og informasjon om det private nettverket blir skjult for omverdenen. En NAT-enhet er som regel kombinert med en brannmur eller en ruter.
Det finnes flere ulike former for NAT. Den enkleste formen er statisk NAT der hver IP-adresse bak en NAT-enhet har en korresponderende IP-adresse på utsiden av NAT-enheten. Dette krever at det er en offentlig IP-adresse per IP-adresse i det private nettverket. En annen tilsvarene variant er dynamisk NAT. Forskjellen er at med dynamisk NAT er det færre IP-adresser på utsiden av NAT-enheten er det er på innsiden. Alle brukerne på innsiden kan derfor ikke ha kontakt med utsiden samtidig. Brukerne på innsiden bytter derfor mellom et gitt antall IP-adresser på utsiden etter et ”første mann til mølla” prinsipp. Den mest vanlige formen for NAT i dag er derimot å sørge for at flere brukere av et privat nettverk kan aksessere et globalt nettverk gjennom en svært få eller en enkelt IP-adresse og visa versa. Denne formen for NAT kalles ofte for Network Address Port Translation(NAPT). I motsetning til statisk og dynamisk NAT benytter NAPT både portnumre og IP-adresser for å oversette adresseinformasjon mellom utsiden og innsiden. Denne oversettelsen lagers i tabeller i NAT-enheten.

I tillegg til disse tre hovedformene for NAT definerer IETF i RFC 3489[43] noen ulike variasjoner av hvordan NAT kan benyttes:

- Fullverdig konus: Her vil en forespørsel fra en bruker på det private nettverket fra en bestemt IP-adresse og et bestemt portnummer alltid oversettes til en bestemt IP-adresse og portnummer i det offentlig nettet. Klienten bak en slik NAT-enhet kan kontaktes utenfra hvis forespørselen kommer på den eksterne adressen.
- Konus med restriksjoner: Her er virkemåten nesten den samme som hos en løsning med fullverdig konus. Forskjellen er at ekstern trafikk kun kan komme inn til en klient bak NAT-enheten fra en hvis klienten først har sendt ut en pakke til den enheten i det eksterne nettverket.
- Symmetrisk: Her dannes det hele tiden nye adressebindinger mellom den interne og den eksterne adressen etter hvert som klienten i det private nettet kontakter nye enheter i det eksterne nettverket.

2.6.3 Hvordan kombinere IP-telefoni og Network Address Translation

Vi har tideligere sett hvordan to brukeragenter utveksler adresseinformasjon under oppsett av en SIP-sesjon. Hvis en eller begge brukeragentene befinner seg i et privatnettverk som benytter NAT vil dette bli et problem. SIP-enheten vil da kun ha kjennskap til adresseinformasjonen på det private nettverket. Det er denne adresseinformasjonen den benytter i SIP-meldingene. Siden NAT-enheten omskriver IP-adresser og/eller portnummer vil den informasjonen brukeragentene utveksler være ugyldig. I dette kapittelet vil det derfor bli presentert noen ulike metoder for å løse dette problemet.

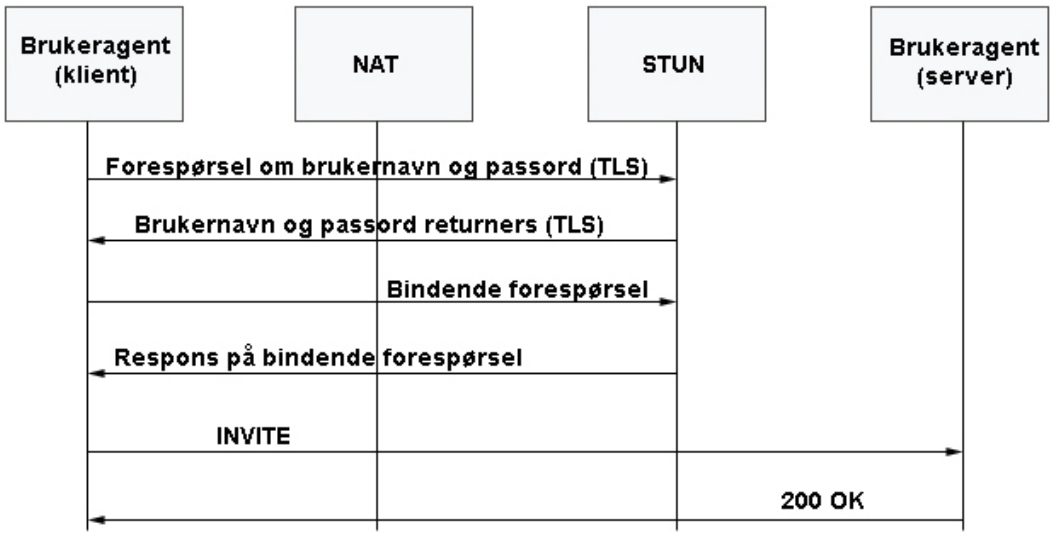
Holde forbindelsen åpen

Når en brukeragent skal benytte en SIP-tjeneste må det først sendes en SIP Register forespørsel fra brukeragenten og til en SIP registerserver som holder oversikt over bindingen mellom brukernavn og brukeren sin IP-adresse. I forbindelse med symmetrisk NAT så vet vi at klienten får tildelt nytt eksternt portnummer for hver forbindelse som settes opp med omverdenen. Portnummeret som ble benyttet da klienten kontaktet SIP-serveren vil altså ikke være gyldig neste gang klienten skal kontaktes, og forsøk på kontakt med brukeragenten vil da bli avvist av NAT-enheten. En løsning på dette problemet er å holde en konstant åpen forbindelse mellom brukeragenten og registerserveren. Hvis det er en UDP forbindelse mellom klient og serveren er klienten avhengig av å sende stadig nye Register forespørsler for å holde forbindelsen åpen. Dette er ikke nødvendig hvis det er en TCP forbindelse. Denne løsning forutsetter at Registerserveren holder oversikt over både IP-adresse og portnummer, samt at brukeragent-klienten befinner seg bak en NAT-enhet. For å løse dette benyttes et tillegg i Via-headerfeltet som kalles ”rport”. Når en klient sender en forespørsel legger den til ”rport” for å bekrefte at den støtter åpne forbindelser. Deretter legger serveren til informasjon om klienten sin eksterne IP-adresse og portnummer i videre kommunikasjon med andre servere. Respons meldinger tilbake til klienten sendes da over den åpne forbindelsen til denne IP-adressen og portnummeret.

Simple Traversal of UDP through NAT(STUN)

Problemer med IP-telefoni og NAT gjelder spesielt overføringen av selve talen som benytter RTP over UDP. Dette problemet er behandlet i [43]. Her foreslås det en løsning som kalles Simple Traversal of UDP through NAT(STUN). STUN er basert på en klient/server arkitektur. Klienten vil som oftest være en brukeragent som befinner seg bak en NAT-enhet og serveren er en Internett-server. STUN benytter to ulike typer forespørsler. Dette er ”bindende forespørsel” og ”forespørsel om delt hemmelighet”.

Figur 2.6.2 ”Klient bak NAT kommuniserer med en STUN-server” modfisert fra [1]



Når klienten sender en ”bindende forespørsel” til serveren må klienten også autentisere seg med en delt hemmelighet. I praksis vil dette som regel være et brukernavn og et passord. Dette brukernavnet og passordet har klient og serveren på forhånd utvekslet over en transport Layer Security(TLS) forbindelse. En forespørsel fra klienten om å motta brukernavn og passord kalles forespørsel om ”delt hemmelighet”.

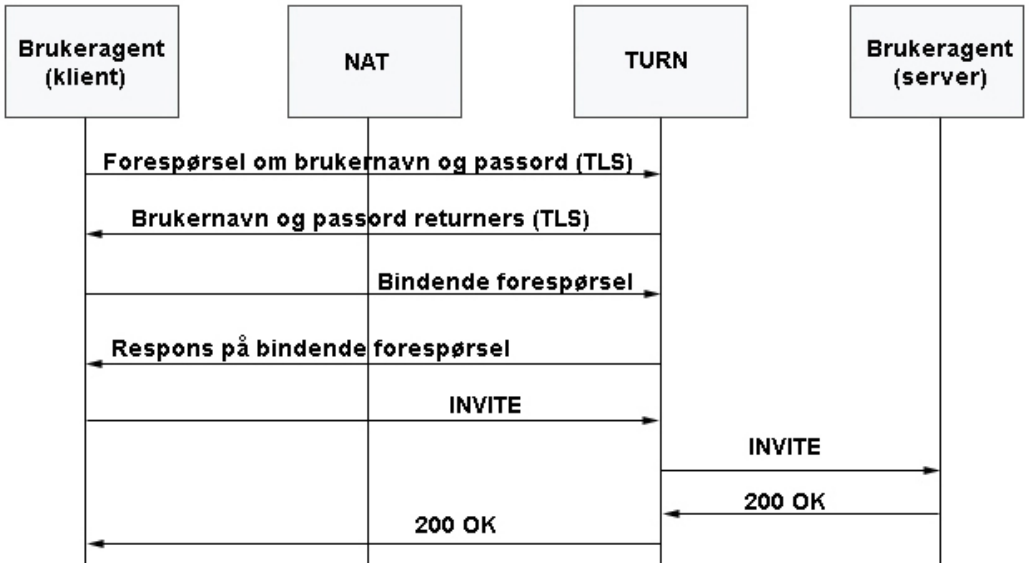
”Bindende forespørsler” sendes fra klienten og ut til serveren over UDP. Formålet med dette er at klienten ønsker informasjon om hvilke IP-adresse og portnummer som er generert av NAT-enheten. Serveren undersøker IP-adressen og portnummeret i forespørselen fra klienten og kopierer dette inn i en responsmelding som sendes tilbake til klienten. Dette vil da være klienten sin eksterne adresseinformasjon. Klienten undersøker så IP-adressen og portnummeret i responsmeldingen med den lokale IP-adressen og

portnumrene som ble benyttet da forespørselen ble sendt. Hvis den lokale og den eksterne adresseinformasjonen ikke stemmer overens vet klienten at den befinner seg bak en NAT-enhet. Men siden klienten nå vet sin eksterne IP-adresse og sin eksterne kildeportnummer kan klienten bruke denne adresseinformasjonen når den skal initiere en SIP-sesjon. STUN har ikke støtte for symmetrisk NAT. Derfor har det kommet et tillegg til STUN som kalles Traversal using Relay NAT(TURN)

Traversal using Relay NAT(TURN)

Som sagt fungerer ikke STUN for alle løsninger som benytter NAT. I en del tilfeller trenger derfor klienten en server på utsiden av NAT-enheten som kan handle på vegne av klienten. Traversal using Relay NAT(TURN) er et eksempel på en slik server.

Figur 2.6.3 "Klient bak NAT kommuniserer med en TURN-server" modifsert fra [1]



I motsetning til en STUN-server ser vi her at en TURN-server videresender trafikk på vegne av klienten. I figur 2.6.2 ser vi derimot at denne kommunikasjonen foregår direkte mellom brukeragentene. Siden alle meldingene til og fra klienten er nødt til å gå innom en TURN-server vil denne serveren bli en flaskehals i nettverket og faren for forsinkelser og pakketap øker. TURN bør derfor ikke benyttes dersom det kan unngås.

Symmetrisk Real Time Protocol

Som vist er problemet at en NAT-enheten omskriver adresseinformasjonen som kommer fra brukeragent-klienten og at portnumrerne for RTP-strømmen ikke lenger er gyldig. En løsning på dette problemet er at begge brukeragentene sender RTP-trafikken på samme portnummer isteden for at de begge velger et tilfeldig portnummer.

I et vanlig oppsett vil som kjent begge brukeragentene først utveksle portnumre i SDP og deretter starte mediaoverføringen. Dette prinsippet fravikes ved bruken av symmetrisk RTP. Her vil først brukeragent(klient) sende RTP-pakker til brukeragent(server). Deretter vil brukeragent(server) benytte det samme portnummeret til å sende trafikk tilbake. Portnummeret som er beskrevet i SPD vil dermed ignoreres og portnummeret som trafikken kommer på vil benyttes isteden.

Et slikt oppsett må initieres av brukeragent(klient). Dette gjøres ved å legge til et ekstra felt i SDP-meldingen *a=direction acitve.* Brukeragent server viser at den støtter tillegget ved å legge til *a=direction passive.*

Kombinasjon av ulike metoder

Som forklart så finnes det flere ulike metoder for å kombinere NAT og SIP. Ofte må disse metodene kombineres med hverandre avhengig av hvilke løsninger som benyttes. Interactive Connectivity Establishment[44] er et rammeverk som beskriver hvordan disse ulike metodene kan kombineres utfra ulike forutsetninger, blant annet ved å definere nye tillegg som kan benyttes i SDP. I kapittel 2.8 vil det bli gjennomgått et eksempel om hvordan denne problematikken kan løses ved bruk av Session Border Controllers.

2.7 Telefonnumre, IP-adresser og ruting

I forbindelse med IP-telefoni er det en del problemstillinger i forbindelse med bindingen mellom IP-adresser og telefonnummer.

- Gitt at en IP-telefonisamtale skal terminere i PSTN-nettet. Problemet er da å finne IP-adressen til en gateway som er i stand til å sette opp denne samtalen.
- Gitt at du vet PSTN-nummeret til en bruker. Samtidig ønsker du å oppdrive IP-adressen til den samme brukeren.
- Hvordan PSTN signalering kan transporteres over et IP-nett.

I dette kapittelet vil noen ulike metoder for å løse disse problemene bli gjennomgått.

2.7.1 Telephony Routing over IP(TRIP)

Telefonruting over IP(TRIP) er beskrevet i RFC2871[45] fra IETF. Av de tre problemstillingene som ble ovenfor er TRIP konstruert for å håndtere den første. TRIP benyttes i hovedsak av tjenestetilbydere for å utveksle rutingtabeller for sine respektive gatewayer. Hvilke gateway som skal benyttes er avhengig av flere ulike forhold, både tekniske og administrative. Eksempler på forutsetninger for å velge en gateway kan være:

- Hvilke former for signalering som støttes av en aktuell gateway.
- Hvor stor kapasitet har hver enkelt gateway.
- Hvilke økonomiske og tekniske samtrafikkavtaler som gjelder mellom tilbyderne som kontrollerer hver enkelt gateway.

Fra disse punktene ser en at samkjøring mellom de ulike tilbyderne kan bli en ganske kompleks affære. Hvilken rutinginformasjon hver enkelt tilbyder vil bidra med vil kort sagt være avhengig av hvem som spør. Konklusjonen til IETF at det derfor er urealistisk å implementere denne informasjonen i en global database. De ulike tilbyderne er derfor avhengig av å utveksle gateway rutinginformasjon internt. TRIP ble konstruert som et rammeverk for hvordan dette kunne gjøres.

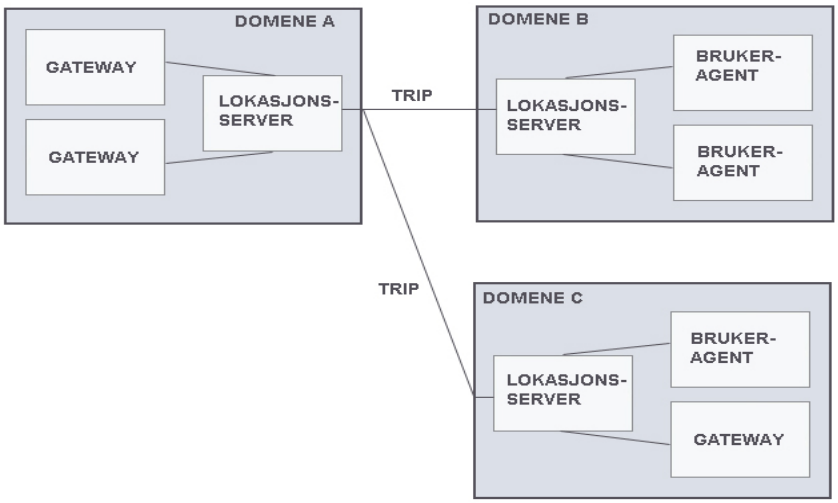
TRIP er en interdomene rutingprotokoll for gatways. Den har arvet en del av virkemåten til Border Gateway Protocol(BGP) som er interdomene rutingprotokoll for IP-trafikk. Allikevel er det en del ting som skiller TRIP fra BGP.

- TRIP befinner seg på applikasjonslaget, i motsetning til BGP som er på nettverkslaget. Informasjonen som utveksles i TRIP er derfor mer komplisert.
- TRIP kan benyttes mellom servere som befinner seg på to ulike nettverk, uten direkte kontakt.
- TRIP binder sammen flere små administrative områder. BGP søker å binde samme flere nettverk til et globalt sammenhengende nettverk.

Arkitektur

Før gateway rutinginformasjon kan utveksles må det på forhånd bestemmes hvilke domener som skal kunne kommunisere med hverandre. Dette er en administrativ avgjørelse som taes uavhengig av TRIP. Etter at dette er avklart kan de aktuelle lokasjonsserverne i hvert domene utveksle gateway rutinginformasjon med hverandre.

Figur 2.7.1 "Overordnet arkitektur av domener som benytter TRIP" modifisert fra [45]



I figuren er det tatt med tre ulike typer domener. Domene A har ikke egne brukeragenter, men tilbyr gatewayfunksjonalitet til andre domener. Vi ser at domene B ikke har noen egen gateway. Ved å benytte TRIP for å utveksle gateway informasjon med andre domene vil domene B allikevel kunne tilby brukerne sine mulighet for å ringe ut til PSTN-nettet. Domene C har både egen gateway og egne brukeragenter.

Gateway

Den grunnleggende virkemåten til en gateway ble omtalt i kapittel 2.5. TRIP en metode å utveksle gatewayinformasjon mellom ulike domener. Denne informasjonen utveksles ved hjelp av TRIP-objekter. Attributtene i et TRIP-objekt baseres på parametere som lokasjonsserveren har mottatt fra sine lokale gatewayer. Eksempler på attributter er:

- Hvilket nummerområde i PSTN-nettet som kan nåes fra en gateway. Dette kan være en stor blokk eller flere mindre blokker. IETF foreslår i tilegg at det presenters en form for kostnadsparameter for hvert enkelt nummerområde.
- Hvor mange samtidige brukere som kan håndteres.
- IP-adresseinformasjon for å nå inn til en gateway.
- Båndbredde inn til en gateway.
- Hvilke signaleringsprotokoller som kan benyttes.
- Hvilke ulike former for talekoding som støttes.
- Hvilke former for krypteringsalgoritmer som støttes.

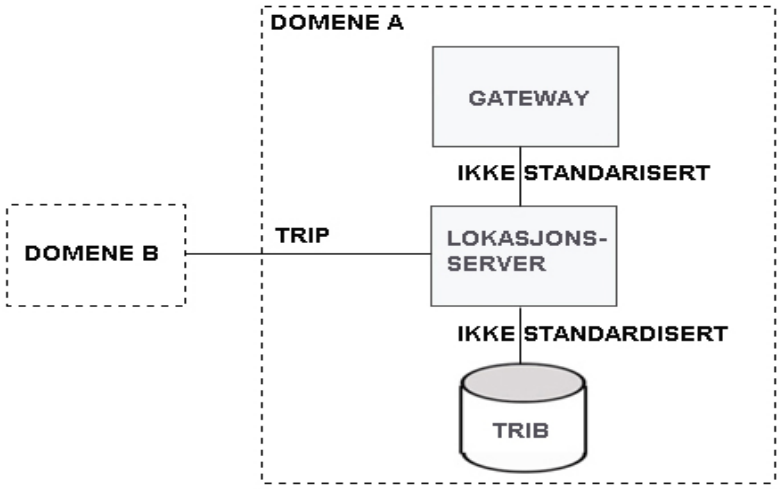
Det er ikke et krav at et TRIP-objekt skal inneholde alle disse attributtene. Et viktig element innenfor TRIP er aggregering, det vil si at flere objekter skal kunne slåes sammen til et objekt. Jo færre tilleggsparametre som tilknyttes hver enkelt gateway, jo enklere er det å slå sammen objektene. Hvis et domene A har utvekslet gateway informasjon med annet domene C kan A presentere sin gatewayinformasjon på to måter ovenfor domene B:

1) Som to objekter: Da må det oversendes et objekt for hvilke telefonnumre som kan nåes fra domene A og et objekt for hvilke telefonnumre som kan nåes fra domene C. I tillegg hvilke parametere som gjelder for de to domenene.
2) Som et sammenslått objekt: Domene A slår sammen telfonnumrenene som kan nåes fra både A og C. Denne informasjonen presenters ovenfor B som et objekt som forteller at alle disse numrene kan nåes via domene A.

Kommunikasjon mellom de ulike enhetene innad i et domene

Som sagt er TRIP en standard for å utveksle informasjon mellom domener. Hva slags standarder som skal brukes til å utveksle informasjon innad i domenet er ikke tatt stilling til.

Figur 2.7.1.2 ”Oversikt over hvor TRIP benyttes”



Lokasjonsserver

Som vist i 2.7.1 har hvert domene en egen lokasjonsserver som benyttes til å utveksle TRIP-informasjon med andre domener. Gatewayinformasjon lagres i en Telefonruting informasjonsdatabase(TRIB). Informasjonen i denne databasen er informasjon om de lokale og eksterne gatewayer som lokasjonsserveren har mottatt fra andre domener ved hjelp av TRIP.

Det mest vanlige bruksområdet for informasjonen som er tilgjengelig i lokasjonsserveren er at den aksesseres av signaliseringsserverne som skal videreformidle forespørsler på vegne av brukeragenter. En lokasjonsserver og en signaleringsserver kan også implementeres som en enhet.

2.7.2 E.164 Number Mapping(ENUM)

Enum er en protokoll fra Internett Enigineering Taskforce(IETF). Bakgrunnen for Enum var at brukerne av skulle kunne benytte sine tradisjonelle E.164 telefonnumre som identifikasjon også innenfor IP-nettet. ENUM beskriver hvordan en kan benytte ”The Domain Name System”(DNS) for å lagre informasjon om E.164-numre. Et eksempel på bruk av ENUM er i samtrafikk mellom en bruker av PSTN-nettet og en bruker av IP-telefoni. I forbindelse med ENUM benyttes det tre ulike typer DNS-oppslag:

Naming Authority Pointers(NAPTR)

NAPTR benyttes for å identifisere mulige måter å kontakte en node på utfra hvilke tjenester som etterspørres.

For å få til en mapping mellom E.164-numre og DNS er det en del utfordringer som må løses. En internettadresse, for eksempel ”www.eksempel.no”, har den mest generelle delen til høyre. Et E.164-telefonnummer, for eksempel +47 73 59 50 00, er bygd opp på motsatt måte. Her vil landskoden ligge lengst til venstre i nummeret. For å gjøre E.164 mer kompatibelt for oppslag i DNS har en derfor i utviklingen av ENUM endret rekkefølgen på E.164-nummeret når det skal gjøres oppslag i DNS. Hvert siffer skilles med et punktum og deretter legges det til ”e164.arpa” for å gjøre DNS søkestrengen komplett. Nummeret +47 73 59 50 00 vil for eksempel formuleres ”0.0.0.5.9.5.3.7.7.4.e164.arpa” som en DNS-søkestreng. Et svar på denne søkestrengen vil kunne se slik ut:

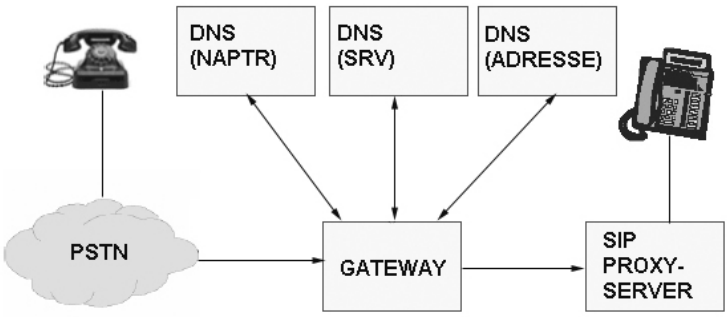
$ORIGIN 0.0.0.5.9.5.3.7.7.4.e164.arpa
IN NAPTR 100 10 "u" "E2U+sip" "!^.$!sip:testbruker@eksempel.no!" .*
IN NAPTR 102 10 "u" "E2U+mailto" "!^.$!mailto:testbruker@eksempel.no!" .*

Verdien 100 forteller oss at brukeren ønsker å kontaktes på adressen *sip:tesbruker@eksempel.no* framfor *mailto:testbruker@eksempel.no* som har verdien 102. Flere linjer med adresseinformasjon kan altså benyttes hvis brukeren har flere måter han ønskes å kontaktes på.

DNS Service Records(SRV)

Ved å benytte SRV kan en klient forespørre DNS om en bestemt protokoll eller tjeneste for et bestemt domene Dette er nyttig i forbindelse med SIP fordi selv om SIP-URI er kjent må det likevel gjøres en forespørsel mot DNS for å finne ut den aktuelle proxyserveren for domenet som brukeren tilhører. Hvis en skal finne hvilken SIP-proxyserver som skal benyttes for å kontakte brukeren testbruker@eksempel.no kan en altså gjøre følgende oppslag _sip._tcp.eksempel.no. Svaret på forespørselen vil da inneholde domenenavnet på den aktuelle SIP-proxyserveren for dette domenet. Samtidig gir SRV records muligheter for å registrere mange proxyservere for hvert domene slik at trafikken mellom de ulike serverne kan balanseres. Hvis en server er opptatt kan neste server på listen kontaktes og så videre.

Figur 2.7.2 ”Telefonisamtale fra PTSN til SIP ved hjelp av ENUM”

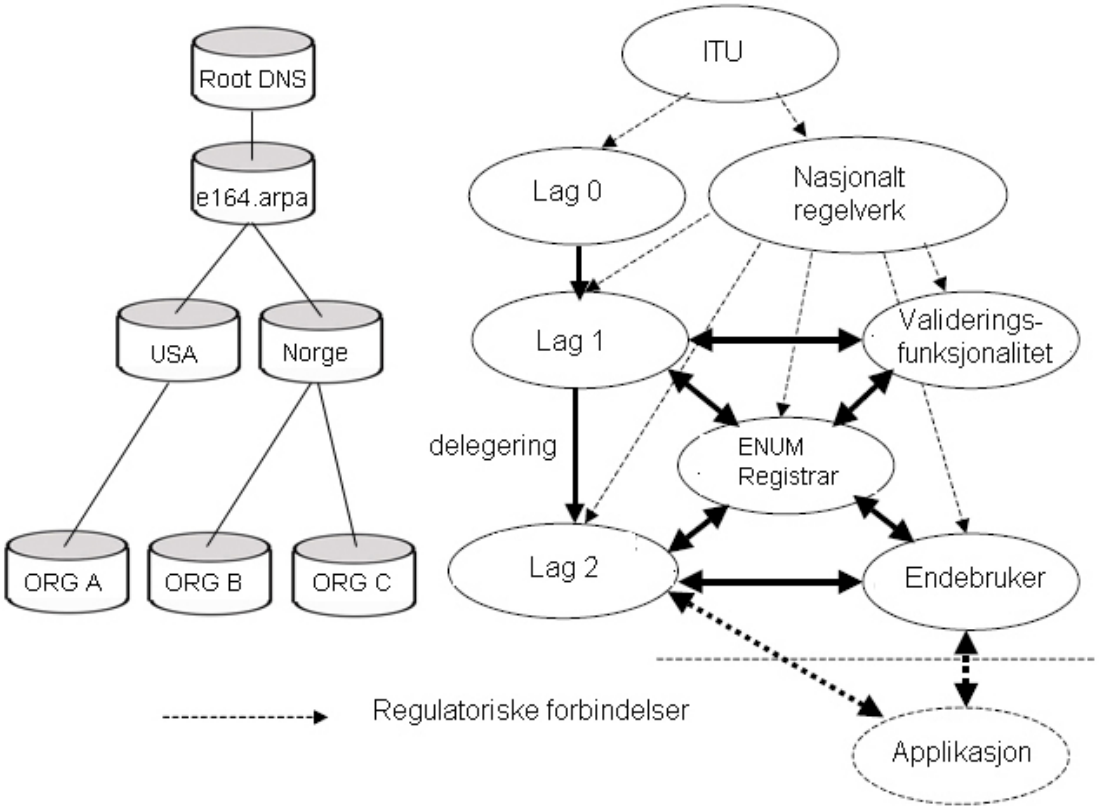


1) Bruker A i PSTN-nettet ringer opp bruker B ved å taste E.164 nummeret.
2) Denne forespørselen rutes så videre til en gateway.
3) Gatewayen må så gjøre om E.164 nummeret til en gyldig DNS-søkestreng. Nummeret +47 73 59 50 00 vil formuleres ”0.0.0.5.9.5.3.7.7.4.e164.arpa” som en DNS-søkestreng.
4) DNS returnerer så respons i form NAPTR. Responsen vil være av typen ”sip:testbruker@eksempel.no”.
5) Gateway må deretter gjør et SRV oppslag for å hvilken SIP-server(e) eksempel.no benytter.
6) Deretter gjøres et vanlig DNS adresseoppslag for å finne IP-adressen til SIP-serveren. Tilslutt videresendes forespørselen til proxyserveren.

Registreringsprosses i DNS i forbindelse med ENUM

De ulike tjenestetilbyderne har et ønske om at ENUM registrering skal realiseres ved å benytte en registerstruktur delt opp tre lag. Et registrar er ansvarlig for å publisere og distribuere ulike sonefiler ut til navneservere for bruk i DNS.

Figur 2.7.3 "Lagdelt struktur i forbindelse med ENUM" modifisert fra [24]



Lag 0: Lag 0 er ansvarlig for domenet *e164.arpa*. Lag 0 inneholder pekere til de ulike landskodene som representer de ulike landene.

Lag 1: Lag 1 er ansvarlig for de ulike E.164 landskodene. Lag 1 inneholder pekere til hvilke domener som har ansvaret for de ulike E.164-numrene

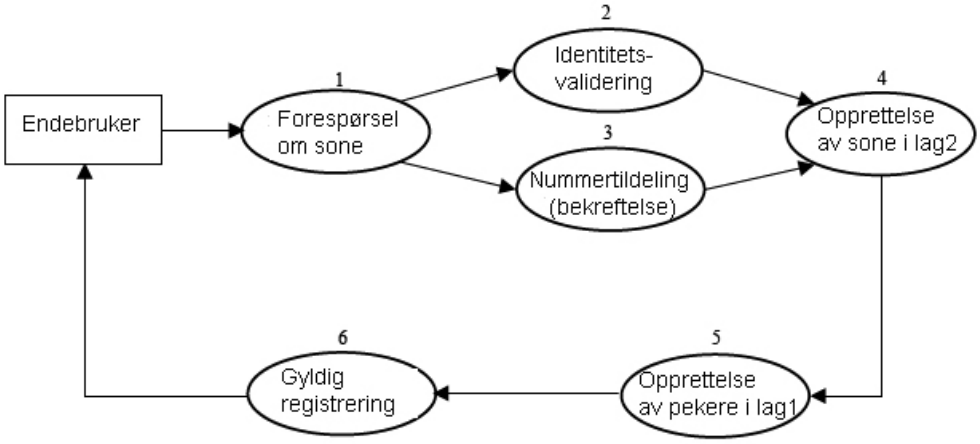
Lag 2: Lag 2 representerer selve E.164 nummeret.

Her lagres NAPTR-data for hvert enkelt telefonnummer.

Registrar: Registraren er forbindelsen mellom endebrukeren og registreringsautoriteten. Et registrar blir dermed kontaktpunktet utad mot endebrukerne.

Endebruker: Dette kan være eieren av nummeret eller en tredjepart som handler på vegne av eieren.

Figur 2.7.4 "Registreringsprosses i DNS i forbindelse med ENUM" modifisert fra [24]



1) Det første som skjer er at en endebruker sender en forespørsel om å danne en DNS-sone for et E.164 nummer. En slik forespørsel kan for eksempel være rettet mot en webserver hos et registrar.
2) Her valideres identiteten til endebrukeren mot et valideringssytem. Det må gis en bekreftelse på at endebrukeren har de nødvendige rettigheter for å foreta endringen.
3) Etter at valideringen er fullført gis det en bekreftelse på at nummerregistreringen kan starte.
4) Det første som skjer i nummerregistreringen er at det opprettes en sone i lag 2 med NAPTR resource records for det aktuelle E.164 nummeret.
5) Etter at det er opprettet en sone i lag2 må det legges inn pekere i lag 1. Lag 1 holder dermed oversikt over hvilke domener som har ansvaret for de ulike numrene.
6) Det gis tilslutt en tilbakemelding til endebrukeren om at fullstendig registrering er gjennomført.

Ulike sikkerhetsmekanismer i forbindelse med en slik arkitektur er nærmere beskrevet i kapittel 7.1.2.

Session Peering for Multimedia Interconnect (Speermint)

Forkortelsen Speermint er en beskrivelse av en arbeidsgruppe som jobber med problemstillinger i forbindelse med samkjøring av ulike nettverk(peering)[13]. Det jobbes imot samtrafikk mellom sanntidsapplikasjoner. Fokuset er derfor på applikasjonslaget og ikke på IP-laget. Arbeidsgruppen som jobber med ENUM er i hovedsak opptatt av spørsmål i forbindelse med oversettelsen mellom E.164-numre og SIP-URI. Speermint arbeidsgruppen jobber med spørsmål i forbindelse med bruk av SIP-URI for å sikre samtrafikk. Det vil si problemstillinger i fasen etter at nummeret er oversatt til en SIP-URI og inntil forespørselen ankommer ingresspunktet på det aktuelle domenet.

Foreløpig har arbeidsgruppen bare gitt ut et utkast som kom februar 2006[35]. Fra dette utkastet gis det en del ulike krav som bør tilstrebes for å realisere fullverdig samtrafikk:

1) Speermint skal være fleksibelt. Det vil si at Speermint skal kunne adopteres både av dagens systemer og fremtidige systemer.
2) Selv om ENUM introduserer oversettelse mellom E.164 og SIP-URI skal fortsatt SIP-URI være den ledende adressestandarden for ruting. Det vil si at domenet som inngår i en SIP URI alltid vil ha hovedansvaret for denne adressen.
3) Systemet skal være skalerbart og fungere uavhengig av hvordan de underliggende nettverkene er bygd opp. Hvorvidt det for eksempel benyttes NAT blant de ulike tilbyderne er en problemstilling som må løses av hver enkelt tilbyder og skal løses av Speermint.
4) Det skal legges til rette for at grunnlaget for om to tilbydere vil ha samtrafikk kan baseres enten på et teknisk eller administrativt grunnlag.
5) Speermint skal ikke forårsake ekstra kostnader eller forsinkelser. Speermint skal også i teorien kunne benyttes av andre protokoller en SIP.

2.7.3 THE INTERNET 2 SIP.EDU INITIATIVE

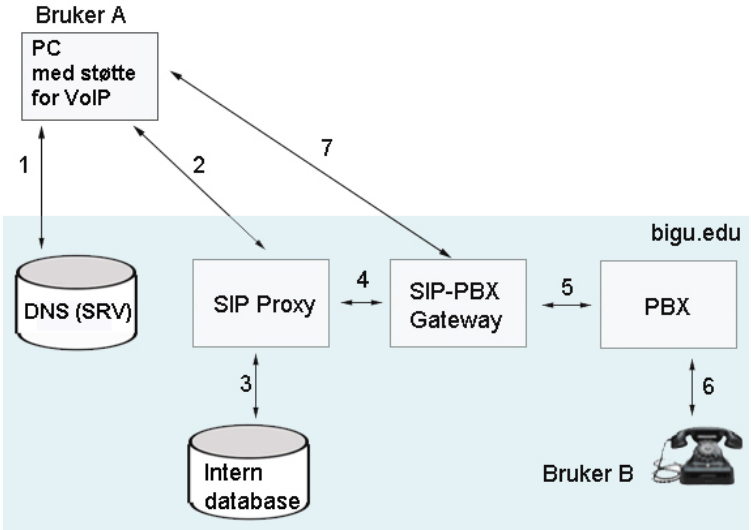
Prosjektet *University Corporation for Advanced Internet Development* (UCAID) beskrives også som Internet2. Prosjektet er et samarbeid mellom over 200 ulike universiteter, andre nettverkespartnere og teknologiselskaper som blant annet Intel og Sun. SIP.edu er en av Internet2 sine mange prosjekter. Målet er å tilby tjenester som VoIP, video og instant messaging gjennom åpne tjenester og standarder. For å realisere dette benyttes SIP som et sentralt verktøy. Dette kapitelet viser to eksempler på SIP.EDU basert på informasjon fra hjemmesiden til The Internet2 SIP.edu Initiative[38].

Samtaleoppsett mot tradisjonelle PBX systemer

SIP.edu benytter eksisterende adressearkitektur og brukernavn for å tilby sine brukere en telefonitjeneste. E-postadressen til brukerne kan enkelt gjøres om til SIP-adresser og benyttes som kontaktinformasjon. For å realisere en IP-telefonitjeneste må i tilegg den eksisterende infrastrukturen utvides med:

- En DNS Service Record tjeneste for å holde oversikt over gyldig proxyserver(e).
- En SIP proxyserver med tilgang på en database med oversikt over oversettelsen mellom brukernavn og interne telefonnumre.
- En gateway for å sørge for samtrafikk med det tradisjonelle telefonisystemet.

Figur 2.7.5 "Standard oppsett av samtale i en SIP.edu arkitektur" modifisert fra [38]



1) Etter at bruker A(Alice) har tastet inn bruker B(Bob) sin adresse, for eksempel *bob@bigu.edu*, må det gjøres et oppslag i DNS SRV for å finne den Proxyserveren som tilhører Bob sitt domene. Hvis UDP benyttes vil forespørselen bli: *sip.udp.bigu.edu.* Responsen blir deretter adressen til proxyserveren.

2) Etter at Alice nå vet hvilken proxyserver hun skal kontakte sender hun *Invite sip:bob@bigu.edu* til proxyserveren.

3) Proxyserveren gjør så et oppslag i en intern database for å oversette *bob@bigu.edu* til et internt telefonnummer.

4) Det nye adressen blir da *sip:12345@gw.bigu.edu*. Dermed kan proxyserveren sende en Invite meldingen videre til Gateway.

5) Med nummeret 12345 kan nå Gateway kontakte Bob sin telefon via en PBX.

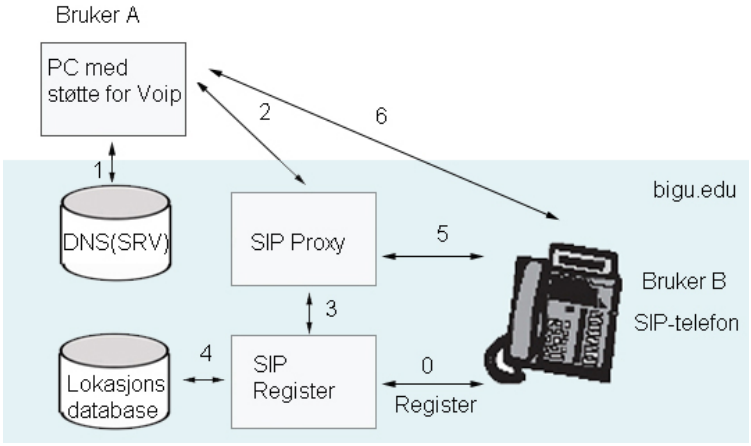
6) PBX kobler forespørselen videre til Bob sin telefon.

7) Etter at samtalen er satt opp går taletrafikken dirkete mellom Gateway og bruker A.

Samtaleoppsett der endebruker benytter SIP-telefon

Etterhvert har SIP.edu i større grad tatt i bruk muligheten for at brukere kan registrere seg med sine respektive SIP brukeragenter. Samtaler til disse brukerne trenger derfor ikke lenger å rutes gjennom en Gateway og PBX. Dette forutsetter at det benyttes en SIP registerserver med en tilhørende lokasjonsdatabase. I det foregående eksemplet så vi hvordan en bruker i det interne telefonnettet kunne kontaktes. Det neste eksemplet vil begge brukere benytte IP-telefoni.

Figur 2.7.6 "Utvidet oppsett av samtale i en SIP.edu arkitektur" modifisert fra [38]



0) Før et slikt system kan benyttes er en avhengig av å ha en registerserver og en lokasjonsdatabase der de enkelte SIP-brukeragentene kan registrere seg.

1) Etter at bruker A(Alice) har tastet inn bruker B(Bob) sin adresse, for eksempel *bob@bigu.edu*, må det gjøres et oppslag i DNS SRV for å finne den Proxyserveren som tilhører Bob sitt domene.

2) Etter at Alice nå vet hvilken Proxyserver hun skal kontakte sender *hun Invite sip:bob@bigu.edu* til Bob sin Proxyserver.

3) Isteden for å oversette brukeradressen til et telefonnummer som i forrige eksempel kontakter proxyserveren nå heller en registerserver for å finne ut om det er registrert en SIP brukeragent for Bob sin adresse.

4) Registerserven sjekker sin database for å finne ut om det er registrert en brukeragent for Bob

5) Hvis dette er tilfelle kan proxyserveren sende en Invite meldning dirkete til Bob sin telefon.

6) Etter at samtalen er satt opp går taletrafikken mellom bruker A og B uavhengig av nettverkselementene som blir benyttet i forbindelse med signaleringen.

2.7.4 SIP for Telephones(SIP-T)

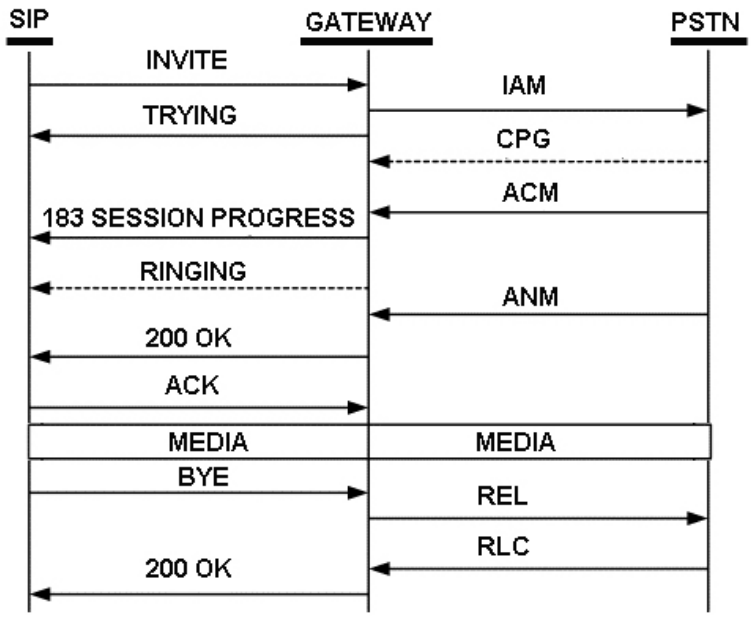
SIP for Telephones(SIP-T) er en beskrivelse av et sett med mekanismer som benyttes for å muliggjøre samtrafikk mellom en tradisjonell telefon i et PSTN-nettverk og en SIP-telefon i et IP-nettverk. SIP-T benytter seg av to ulike teknikker for å realisere dette, oversettelse(mapping) og innpakking(encapsulation). Ved oversettelse blir så mye felles informasjon som mulig oversatt mellom de ulike protokollene. Den viktigste informasjon som må oversettes er adresseinformasjonen. Oversettelse av denne vil bli nærmere omtalt senere. Siden forskjellige protokoller er bygd opp på ulike måter, vil det ikke være mulig å oversette all informasjonen. Løsningen på dette er som oftest å benytte ”default” verdier på de feltene som ikke lar seg oversette direkte. PSTN-signalering er noe mer kompleks enn i SIP, noe informasjon vil derfor gå tapt i oversettelse fra PSTN til SIP. Hvis det er nødvendig å bevare alle elementene fra PSTN-signaleringen er det ikke tilstrekkelig med kun oversetting. I disse tilfellene benyttes derfor innpakking i tillegg til oversettelse. Først oversettes den nødvendige informasjon fra PSTN til SIP, slik at meldingene kan rutes i IP-nettverket. Deretter pakkes PSTN protokollinformasjonen som et vedlegg i SIP meldingen i form av SDP. Slike meldinger vil være i stand til å frakte PSTN informasjon over IP-nettet uten at informasjon går tapt. IP-nettet blir derfor transparent for PSTN-signalering. Forutsetningen for at dette skal fungere er at det benyttes samme type PSTN-signalering i begge ender. I RFC 3372 [32] beskrives det tre ulike scenarioer for hvordan en samtale kan settes opp mellom et PSTN-nettverk og et IP-nettverk som benytter SIP

SIP Oversettelse

PSTN->IP: Samtaler som initieres i PSTN-nettverket traverereserer gjennom en gateway, for deretter å terminere i en IP-telefon i et SIP endepunkt. Siden samtalen her skal terminere i en IP-telefon er det ikke nødvendig å ta vare på alle feltene i PSTN-signaleringen, siden IP-telefonen ikke vil kunne benytte seg av dette uansett. I disse tilfellene er det derfor tilstrekkelig med kun oversettelse, innpakking av PSTN-signalering i SDP er ikke nødvendig.

IP->PSTN: Samtalene som initieres i SIP-nettverket travereserer gjennom en gateway, for deretter å terminere i en PSTN-telefon. Dette tilfellet blir mye av det samme som det første eksemplet. Også her er det tilstrekkelig med kun mapping mellom SIP og PSTN.

Figur 2.7.7 ”Oppsett av sesjon ved hjelp av SIP og ISUP” modifisert fra[30]



SIP INNPAKKING

PSTN->IP->PSTN: IP-nettverket benytter SIP og fungerer som et bindeledd mellom ulike gatewayer. En samtale vil dermed kunne initieres og terminers i PSTN, men i midten vil det være et IP-nettverk. Denne varianten omtales ofte som SIP innpakking. Hvis IP-nettverket i midten skal oppleves som transparent for brukerne av PSTN-telefoner i begge ender er det nødvendig både med oversettelse og bevaring av PSTN-signaleringsinformasjon. PSTN-signaleringen må derfor fraktes i SIP meldingene ved hjelp av SDP.

2.8 Session Border Controllers

I forbindelse med IP-telefoni fungerer Session Border Controller(SBC) som en enhet som stort sett er plassert i ytterkanten av nettverk for å håndtere kommunikasjon med andre nettverk. Hovedfunksjonaliteten til en SBC er å løse problemer som er forbundet med å sørge for mediatransport og signalering via brannmurer eller NAT. SBC er kommersielle komplekse løsninger som kan inneholde både gateway- og ulike former for SIP-serverfunksjonaliet og kundeadministrasjon. Ulempen er at dette er dyre løsninger og bryter med prinsipper fra IETF om at SIP skal være en enkel og åpen protokoll. Foreløpig har ikke IETF gitt ut noen RFC i forbindelse med SBC, men det et gitt ut et ”Internet-Draft” [54].

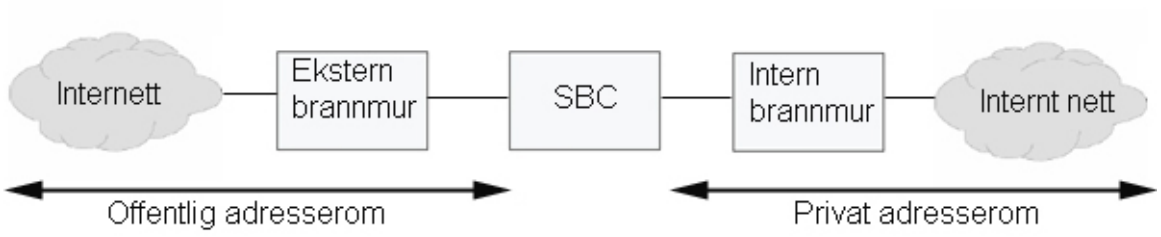
2.8.1 Grunnnleggende arkitektur

En SBC består logisk sett av to deler:

1. SBC Signaleringsfunksjonalitet: Holder oversikt hvilken signalering som skal få aksess til nettverket, og innholdet i denne signaleringen.
2. SBC Mediafunksjonalitet: Holder oversikt over hvilke mediapakker som skal få aksess til nettverket. I tillegg tilbys overvåking av båndbredde, differensierte tjenester og tjenestekvalitet for de ulike mediastrømmene.

Disse to delene kan implementeres som en felles fysisk enhet eller to atskilte fysiske enheter. I det siste tilfellet er det nødvendig å benytte en kontrollprotokoll, for eksempel Megaco som er omtalt i avsnitt 2.6.2, for å kommunisere mellom de to atskilte enhetene.

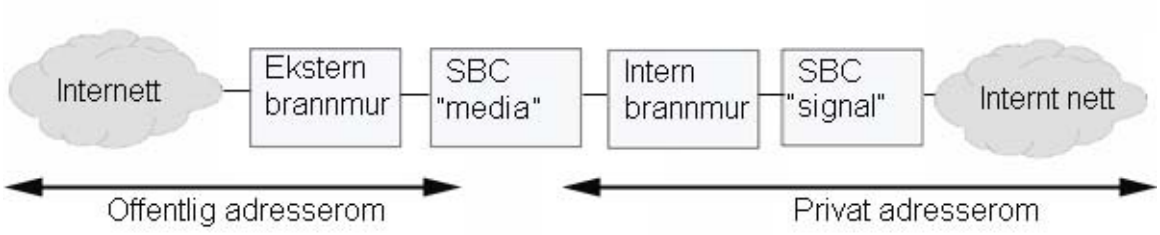
Figur 2.8.1 ”Plassering av enkel SBC mellom to nettverk” modifisert fra [46]



I figur 2.9.1 ser vi et eksempel på en SBC hvor både signalerings og mediafunksjonaliteten er plassert i en og samme boks. Den eksterne brannmuren skal

forhindre at uønsket trafikk kommer inn i det interne nettverket. SBC-enheten kontrollerer signaleringstrafikken og oversetter mellom de offentlige og de private adressene. Den interne brannmuren er satt opp for å kontrollere trafikken ut fra det interne nettverket. Denne brannmuren kan i mange tilfeller utelukkes, men dette avhenger av hvilke regler som gjelder for brukerne av det interne nettverket.
En vanlig konfigurasjon av brannmurene innebærer at disse er statisk programmert til å slippe igjennom all trafikk inn mot SBC-enheten. En annen mulighet er at SBC-enheten selv konfigurer brannmurene dynamisk. Ellers kan både den eksterne og den interne brannmuren også implementeres innad i SBC-enheten.

Figur 2.8.2 ”Plassering av delt SBC mellom to nettverk” modifisert fra [46]



I figur 2.7 ser vi et eksempel på en arkitektur der mediafunksjonaliteten og signaleringsfunksjonaliteten er delt opp i to forskjellige enheter. Ved en slik løsning kan en benytte flere SBC media enheter for å håndtere trafikken, men samtidig benytte kun en sentral SBC sinaleringsenhet.

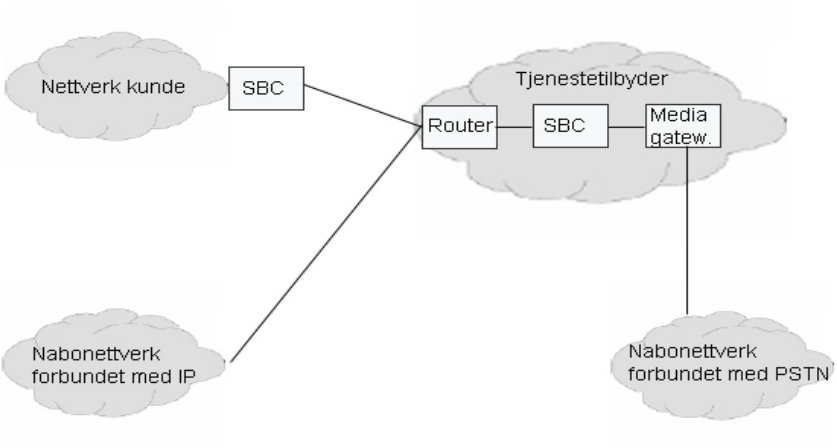
2.8.2 Praktiske bruksområder for Session Border Controllers

Bruksområdene for SBC kan i hovedsak deles opp i fem hovedområder:

1. SBC-enheten plasseres i overgangen mellom nettverket til tjenesetilbyderen og nettverket til kunden.
2. SBC-enheten plasseres i overgangen mellom nettverkene til to tjenestetilbydere for å kontrollere samtrafikk.
3. SBC-enheten benyttes for koordinering av samtaler som foregår over VPN forbindelser.
4. SBC-enheten benyttes for å kontrollere hvor stor del av ett nettverks båndbredde som skal dedikeres til IP-telefoni.

5. SBC-enehten kan benyttes til å oversette mellom ulike IP-telefoner i et nettverk hvis disse er konfigurert til å benytte ulike metoder for koding av tale.

Figur 2.8.3 "SBC mellom tjenestetilbyder og kunde" modifisert fra [46]



Figuren viser et oppsett med SBC-enheter i forbindelse med en tjenestetilbyder som gir et kundenettverk tilgang til IP-telefoni. Samtaler som skal terminere i PSTN-nettverket eller andre nettverk som ikke har tilknytning til tjenestetilbyderen sitt nettverk rutes til PSTN-nettverket via en Media Gateway. Hvis samtalen terminerer i et nettverk som har direkte tilknytning til tjenestetilbyderen sitt nettverk kan denne rutes direkte dit. I denne forbindelsen er det vanlig for tjenestetilbyderen å ha en SBC-enhet mellom kundenettverk og det terminerende nettverket. SBC-enheten håndterer rutingen og denne informasjon holdes derfor skjult for kundenettverket. Samtidig holder SBC-enheten kontroll på områder som samtaleoppsett, forbruk av båndbredde og antall samtaler.

2.8.3 Hvilke funksjoner utføres av Session Border Controllers

Som sagt så finnes det ikke noe standardisert sett av oppgaver som skal utføres av SBC, og ulike SBC har ulik funksjonalitet. Noen av de mest vanlige funksjonene er:

- SBC benyttes til Network Address Translation(NAT) for tjenestetilbyderen.
- SBC fungerer som en brannmur, eller samarbeider med andre brannmurer i nettverket. SBC kan tilby støttefunksjoner for brannmurer ved å tilby filtermuligheter slik at trafikk til IP-telefoni kan passere igjennom

- SBC skjuler tjenestetilbyderen sin interne nettverkstopologi. Som vist i figur 2.8.3 vil et kundenettverk kun forholde seg til en SBC-enhet, og vil ikke ha noe kunnskap om nettverket som ligger bak denne. For å realisere dette er SBC avhengig av å omskrive signaleringsmeldingene. Et eksempel på dette er at ”Via-feltet” i SIP headeren omskrives.
- SBC benyttes til samtalekontroll ved å for eksemepel å overvåke båndbredden som benyttes av hver enkelt bruker.
- SBC kan benyttes i forbindelse med å sikre tjenestekvalitet. Metoder for å tilby tjenestekvalitet er nærmere omtalt i kapittel 7.3.

2.8.4 Bruk av SIP Via i forbindelse med Session Border Controllers

Som vist i avsnitt 2.1 inneholder headerfeltet i en SIP melding et felt som kalles for “Via”. Informasjonen i dette feltet benyttes av mottaker av en melding når en forespørsel skal besvares. Når en brukeragent A sender en Invite forespørsel i forbindelse med oppsett av en SIP-sesjon lagres adresseinformasjon om avsender i “Via-feltet” i SIP-headeren. Etter hvert som denne meldingen transporteres gjennom et nettverk vil hver nye enhet som prosesser meldingen, for eksempel en proxyserver, lagrer sin adresse-informasjon i “Via-feltet”. Den gamle informasjonen i “Via-feltet” vil fortsatt bli liggende, men den nyeste informasjonen legges til på toppen. “Via-feltet” vil dermed bli en liste over hvilke rute SIP-meldingen har tatt gjennom nettverket. Når så meldingen er kommet gjennom nettverket og responsmeldingen skal sendes tilbake kopieres “Via-feltet” fra forespørsel og inn i responsmeldingen.

Figur 2.8.4 ”Eksempel på Via-feltet i SIP” fra [1]

```
Via: SIP/2.0/UDP 100.101.102.103
 ;branch=z9hG4bK776a

Via: SIP/2.0/TCP cube451.office.com:60202
 ;branch=z9hG4bK776a

Via: SIP/2.0/UDP 120.121.122.123
 ;branch= z9hG4bK56a234f3.1
```

Fra figuren ser vi at hver linje inneholder informasjon om hvilken SIP-versjon som benyttes, protokollinformasjon, adresseinformasjon og en cookieverdi. Denne verdien

skal alltid starte med de 7 tegnene “z9hG4bK” og deretter etterfølges den av en kryptografisk hashverdi. Denne hashverdien er beregnet på grunnlag av feltene to, from, CallID og Request-URI i SIP headeren. Ved hjelp av denne informasjonen kan en server som mottar en SIP respons sjekke gyldigheten av denne. Ved å sjekke “Via-feltet” vil altså brukeragent B ha full oversikt over hvilke nettverkselementer forespørslene fra brukeragent A har passert og aktuell adresseinformasjon om disse. Dette står godt i stil med prinsippet om at SIP skal være en åpen protokoll. De ulike tjenestetilbyderne ønsker imidlertid ofte ikke så stor grad av åpenhet i forbindelse med oppsett av samtaler. Ofte benyttes derfor løsninger der SBC-enheten sletter innholdet i Via-feltene, og selv tar ansvaret for rutingen. På denne måten kan tilbyderne skjule informasjon om sitt nettverk for brukerne av IP-telefonitjenesten.

2.8.5 Utfordringer med brannmurer

Brannmurer er en viktig nøkkelfaktor i forbindelse med nettverkssikkerhet. De skal sørge for et trygd skille mellom innsiden av et nettverk og omgivelsene. Innsiden av nettverket skal beskyttes mot autorisert aksess og ulike typer angrep. Et annet bruksområde er å hindre interne brukere tilgang til tjenester på utsiden av brannmuren. Mange brannmurer kombineres nå med NAT funksjonalitet. Det største problemet med brannmurer i forbindelse med IP-telefoni er at de ikke selv er i stand til å prosessere trafikken, noe som er nødvendig for å avgjøre hva som skal få lov til å passere av IP-telefonitrafikk og ikke. Dette problemet blir enda større hvis denne trafikken i tillegg er kryptert.

Ulike virkemåter

Tradisjonelle brannmurer har stort sett filtrert trafikk på grunnlag av IP-adresser og portnumre. Etter hvert er denne formen for pakkefiltrering blitt utilstrekkelig, da avgjørelsen om hvorvidt trafikk skal slippe gjennom brannmuren er avhengig av flere parametre enn kun IP-adresser og portnumre. Det har defor utviklet seg et marked for å implementere Application Layer Gateways(ALG) i samarbeid med brannmurer.
I motsetning til tradisjonell brannmurfunksjonaliet er en ALG istand til å analysere spesifikke deler av innholdet i meldiningskroppen. Denne formen for analyse gjør at ALG er bedre skalert for å håndtere trafikk i forbindelse med SIP og H.323. Denne prosesseringen gjør at det ikke lenger er mulig med direkte trafikk mellom to brukere på

hver sin side av en brannmur, all trafikken må passere gjennom en ALG. ALG regnes derfor som sikrere enn tradisjonelle brannmurer. Ulempen er at en ALG er mer ressurskrevende og dermed også tregere enn en tradisjonell brannmur. Dessuten kreves det at ALG-enheten må konfigureres med et eget sett av regler for hver enkelt applikasjon.

Brannmurer og SIP

I forbindelse med SIP fungerer det bra å kjøre signaleringstrafikken gjennom en brannmur. Siden signaleringstrafikken kan benytte TCP og velkjente portnumre kan brannmuren åpnes for denne trafikken. Problemet oppstår ved overføring av mediatrafikken. Som kjent overføres denne trafikken ved hjelp av RTP-pakker over UDP. RTP-trafikken tildeles dynamiske portnumre i form av partall i intervallet[1024-65534]. Å åpne opp alle disse portene for trafikk vil innebære en sikkerhetsrisko. En mulighet er derfor å implementere brannmurfunksjonalitet i NAT og kombinere dette med en SBC. Dette blir beskrevet i kapittel 2.8.6

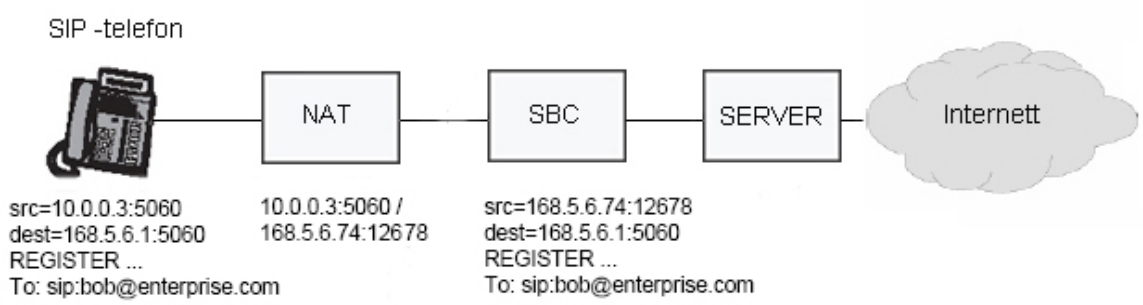
2.8.6 Utfordringer med Network Address Translation og SBC

I denne delen vil det bli gjennomgått hvordan Session Border Controllers fungerer som en mulig løsning på problemene som oppstår i nettverk som kombinerer IP-telefoni med Network Address Translation(NAT). En del ulike løsninger på dette ble omtalt i kapittel 2.7.3. I dette kapittelet vil det bli gjennomgått hvordan dette kan gjøres i forbindelse med Session Border Controllers.

Signalering

SBC kan benytte metoden med å holde forbindelsen åpen som tideligere omtalt. Etter at en forbindelse mellom en SIP-telefon og SBC er opprett kan enten SBC sende meldingen OPTIONS til telefonen eller telefonen sender REGISTER til SBC. Dette er som sagt for å unngå at NAT forbindelsen settes opp på nytt slik at det portnumrene ikke lenger er gyldige.

Figur 2.8.5 ”REGISTER melding med SBC” modifisert fra [46]



I figuren ser vi en SIP-telefon som skal registrere seg hos en tilbyder. Adressen som telefonen skal registre seg hos er *165.5.6.1:5060*. Siden SIP-telefonen befinner seg på et lokalnettverk og ikke vet sin eksterne adresse oppgir den sin lokale adresse *10.0.0.3:5060* som kontaktadresse. Problemet er at SIP-telefonen ikke kan motta meldinger på denne adressen siden den benytter NAT. SIP-telefonen er derfor nødt til å kontaktes på den eksterne adressen 168.5.6.74:12678. SBC registrerer derfor den eksterne adressen i en lokasjonstjeneste på vegne av SIP-telefonen. I tillegg må forbindelsen holdes åpen slik at NAT-enheten ikke genererer en ny adresse på vegne av SIP-telefonen.

Media

Vi har tideligere sett at portnumre i forbindelse med RTP-trafikken beskrives ved hjelp av innholdet i SDP og at det er et problem at portnumre omskrives ved passering gjennom NAT. En vanlig løsning på dette er at SBC omskriver innholdet i SDP på vegne av SIP-telefonen. En slik løsning kan kombineres med symmetrisk RTP. En innkommende RTP strøm til SIP-telefonen sendes derfor først til SBC som oversetter adresseinformasjonen og videresender til den eksterne NAT adressen. Også i dette tilfellet er SBC avhengig av å holde forbindelsen med SIP-telefonen åpen gjennom NAT-enheten hele tiden.

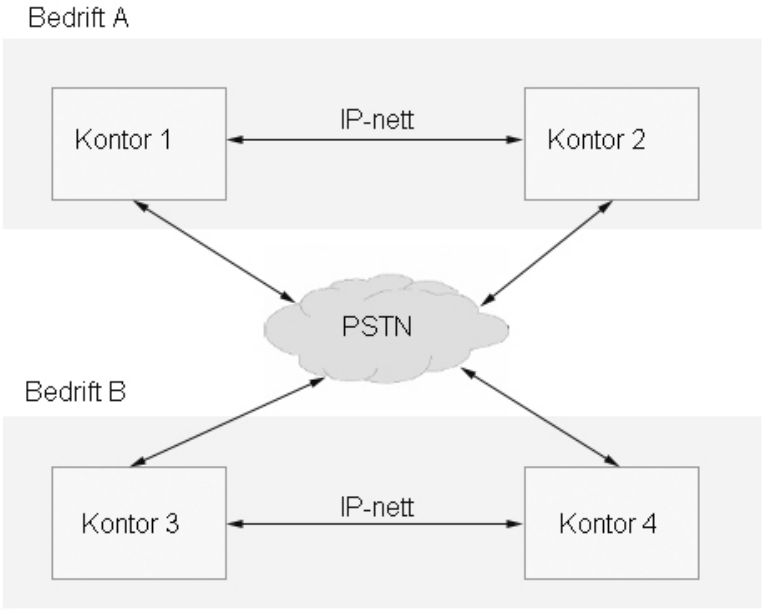
3 Ulike modeller for samtrafikk mellom bedrifter

I dette kapitlet vil det blir gjennomgått fire ulike modeller for hvordan samtrafikk mellom to bedrifter kan løses. Figurene er ment å gi en overordnet oversikt og alle detaljer er derfor ikke tatt med.

Separate løsninger for data og telefoni

En vanlig bedriftsløsning før IP-telefoni ble tatt i bruk var å benytte VPN over et IP-nettverk for å sikre datatrafikk mellom to geografisk spredde kontorer innenfor samme bedrift. Telefonisamtaler mellom kontorene i en bedrift og ut mot andre bedrifter foregikk over PSTN.

Figur 3.2.1 "Atskilte nettverk for tale og data" modifisert fra [21]

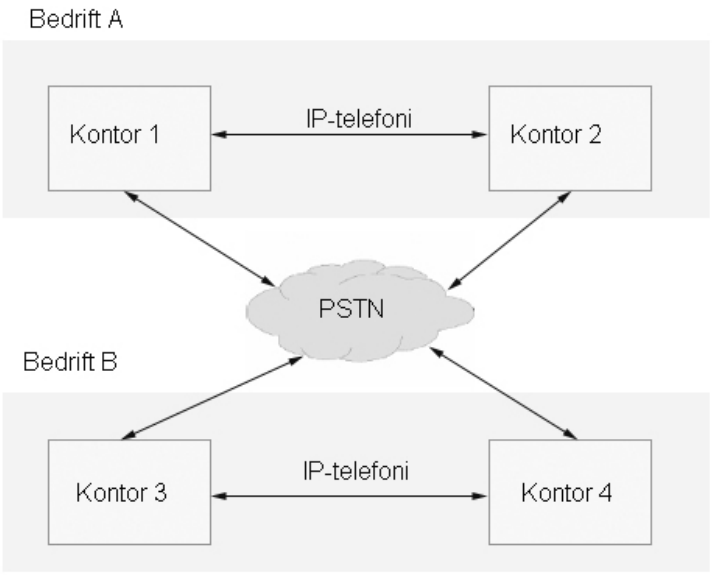


Et slik oppsett gir god sikkerhet og god kvalitet. Ulempen er at det vil være dyrt å benytte PSTN-tjenester for taletrafikk i motsetning til å benytte IP-telefoni.

IP telefoni innad i bedriften

Etter hvert har flere bedrifter begynt å benytte løsninger med IP-telefoni over datanettverket siden de allerede har infrastrukturen på plass. En vanlig løsning på dette er vist i figur 3.2.2.

Figur 3.2.2 ”IP-telefoni for interne samtaler og PSTN for eksterne” modifisert fra [21]

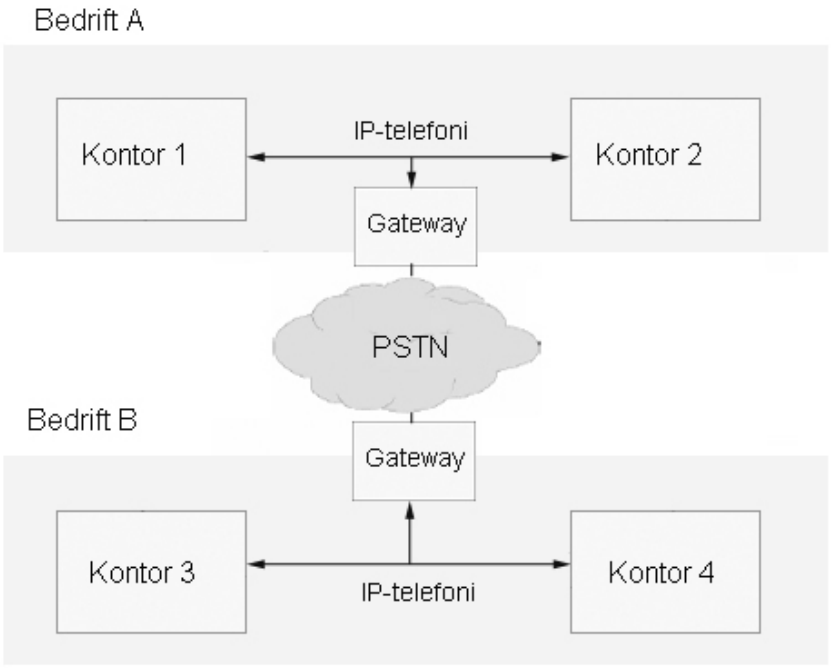


Her foregår trafikk innad i bedriften ved hjelp av IP-telefoni, men kommunikasjon utad til andre bedrifter blir sendt over PSTN-nettet.

Moderne løsninger for IP-telefoni

I disse tilfellene har bedriften gått over til fullverdig IP-telefoni. Det vil si at alle samtaler foregår ved brukt av for eksempel SIP. For samtrafikk med andre bedrifter må allikevel en gateway benyttes for eventuell samtrafikk med PSTN-nettet.

Figur 3.3.3 "Løsninger med bruk av gateway" modifisert fra [21]

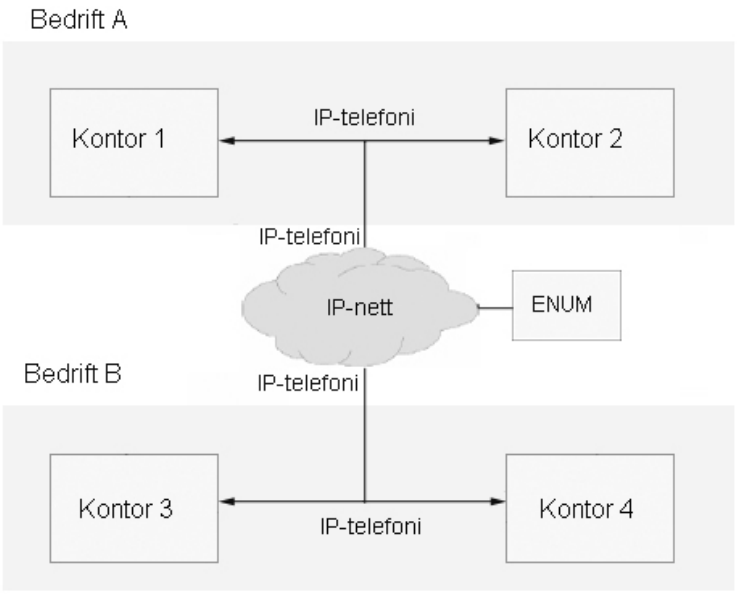


Forskjellen i forhold til det forrige eksempelet er at begge bedriftene har gått over til fullverdig IP-telefoni og har en felles forbindelse ut mot en gateway. For transporten mellom nettverkene ser vi at trafikken kjøres over PSTN-nettet. Ved å benytte TRIP vil en på en mer effektiv måte kunne utveksle gatwayinformasjon for å finne gunstige termingeringspunkter ut mot PSTN. Dermed kan trafikken transporteres mer effektivt og billig siden trafikken ikke trenger å transporters så store avstander over PSTN-nettet.

Fremtidige løsninger for IP-telefoni med bruk av ENUM

Her benytter begge bedriftene løsninger med ENUM. Siden ENUM oversetter E.164 telefonnumre til IP-adresser trenger ikke lenger samtalen å rutes via PSTN.

Figur 3.3.4 "Løsninger med bruk av ENUM" modifisert fra [21]



I figurer ser vi at PSTN-løsninger blir overflødige så lenge begge nettverkene støtter bruk av ENUM. Det er allikevel nødvendig med rutingmuligheter ut mot PSTN hvis det skulle vise seg at det ikke finnes oppslag for enkelte numre i ENUM.

4 Peer-to-Peer løsninger for IP-telefoni over Internett

Peer-to-peer løsninger(P2P) er systemer der all funksjonalitet er lagt ut til de deltakende nodene som deretter samkjøres med hverandre. En blir dermed uavhengig av at kommunikasjonen skal være basert på sentrale servere i nettverket. Slike løsninger er blitt populærere ettersom de er sentrale i en rekke fildelingstjenester. Argumenter som brukes til fordel for et P2P system er at det er mer robust og skalerbart siden systemet ikke er basert på at noen få sentrale servere til enhver tid er tilgjengelige. I dette kapittelet vil det bli gitt en rask gjennomgang av Skype siden dette av mange regnes som en P2P-tjeneste. Tilslutt vil det presenteres hvilke muligheter det er for å realisere en fullverdig P2P-tjeneste ved hjelp av SIP.

4.1 Skype

Skype har over 50 millioner faste brukere og programvaren har blitt lastet ned over 200 millioner ganger. Skype er dermed den mest populære applikasjonen for VoIP over Internett. Det er flere grunner til at denne tjenesten er blitt så populær. Først og fremst skyldes dette at tjenesten er gratis for alle samtaler med andre brukere av Skype, men for samtrafikk med PSTN-nettet må brukerne betale. Dessuten er Skype enkel å laste ned og installere. Tjenesten fungerer fint både i nettverk som benytter NAT og brannmurer, noe som har vært en utfordring for andre former for IP-telefoni. Ellers er en Skype-forbindelse kryptert i motsetning til mange andre IP-telefonitjenester.

Skype regnes ofte som en ren peer-to-peer tjeneste(P2P). Dette stemmer ikke helt da det er en del aspekter ved Skype som gjør den forskjellig fra kravene som stilles til en P2P-tjeneste. For det første benyttes en sentral server for autentisering av brukere. Dessuten benyttes enkelte klientmaskiner som såkalte supernoder. Det vil si at disse enhetene benyttes som viderekoplingspunkter for andre enheter som befinner seg bak en brannmur eller NAT. I tillegg benyttes servere i forbindelse med samtrafikk med PSTN-nettet.

Skype er en lukket løsning og ingen vet helt sikkert hvordan denne tjenesten er oppbygd. Ved å studere virkemåten til Skype har en konkludert med at signaleringen som benyttes verken er SIP eller H323. For kryptering benyttes det, ifølge Skype selv, først RSA i

forbindelse med utveksling nøkler. Deretter benyttes Advanced Encryption Standard(AES) med 256-bits nøkler til å kryptere datastrømmen mellom de ulike klientene.

4.2 SIP som Peer-to-Peer

Det er i utgangspunktektet ingen hindringer i selve oppbyggingen av SIP-meldingene som hindrer SIP fra å bli en fullverdig P2P-arkitektur. Hovedproblemet er at brukerne er avhengige av å lokalisere adresseinformasjon om hverandre. Merk at dette ikke nødvendigvis er det samme som å benytte ENUM Ved å benytte ENUM klarer en bruker som regel å lokalisere den aktuelle proxyserveren for mottakeren av samtalen, mens i P2P er det snakk om å lokaliserere brukeren direkte. A. Johnston og H. Sinnreich foreslår en løsning med bruk av dynamisk DNS i[19], men de påpeker at når funksjonaliteten i en proxyserver er erstattet av en DNS-server kan derfor ikke denne løsningen regnes som noen fullverdig P2P løsning. Dynamisk DNS er nærmere beskrevet i RFC3007[20].

Det er nå flere aktører som har jobbet med utviklingen av SIP som også jobber med krav til hvordan protokollen skal kunne benyttes som en fullverdig P2P protokoll. Det hersker fortsatt en del uenighet om hvordan en slik løsning skal kunne realiseres og det har ikke kommet noen endelige løsninger på dette området fra IETF.

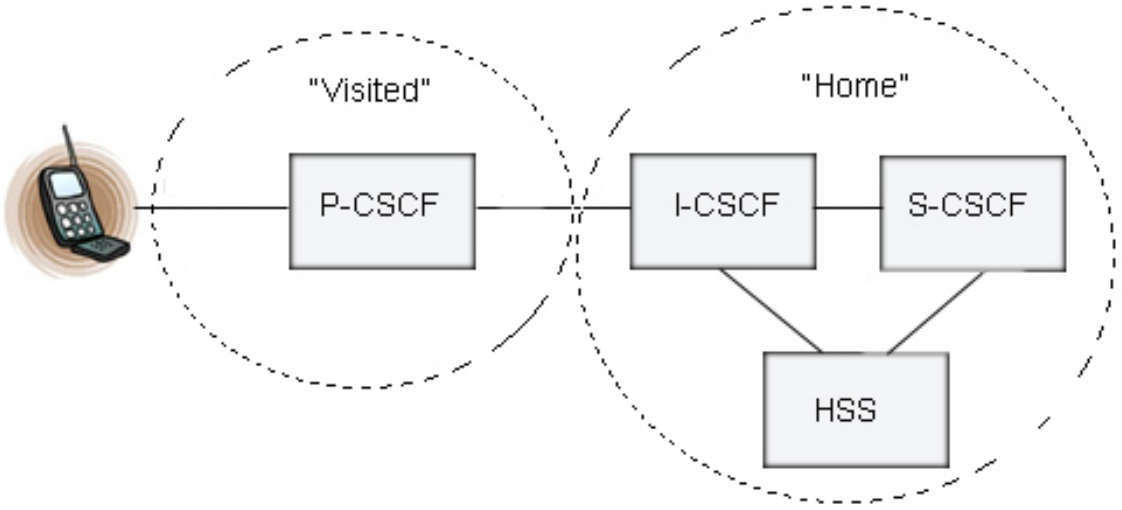
5 The IP Multimedia Subsystem og 3GPP

3GPP er en forkortelse for *3rd Generation Partnership Project.* 3GPP er et samarbeidsprosjekt mellom flere standardiseringsorganisasjoner[1] innenfor telekommunikasjon. En ny spesifikasjon fra 3GPP kalles *release.* Hver *release* bygger på den forrige men inneholder ny funksjonalitet. The IP Multimedia Subsystem(IMS) ble definert i *Release5* og målet er å sørge for en ytterligere sammensmelting av data-, mobile- og linjesvitsjede nettverk. Bakgrunnen for dette er at en ønsker å tilby telefonikunder et bredere spekter av ulike multimediatjenster. Dette realiseres kort sagt ved å benytte SIP på toppen av Universal Telecommunications System(UMTS). Denne løsningen har medført at også IETF har bidratt i standardiseringen av IMS.

5.1 Arkitektur

Det benyttes flere ulike typer av SIP servere for å prosessere SIP signaleringen i IMS. Innen IMS har disse fått navnet Call/Session Control Function(CSCF). Ut fra hvilken funksjonalitet de har, benyttes det tre ulike navn på CSCF serverne;
Proxy CSCF(P-CSCF), Interrogating CSCF(I-CSCF) og Serving CSCF(S-CSCF).
Fullverdig IMS skal benytte IPv6, derfor må dette støttes av alle komponenter.

Figur 5.1 "Sammenheng mellom ulike komponenter i IMS"



[1] ARIB(Japan), CCSA(Kina), ETSI(Europa), ATIS(Nord-Amerika), TTA(Sør-Korea) og TTC(Japan)

Mobil enhet

Mobile enheter i IMS er avhengig av å tildeles et IM Services Identity Module(ISIM)-kort. Dette kortet har en tilsvarende funksjonalitet som SIM i GSM/GPRS og USIM i UMTS. Etter hvert vil IMS også støtte bruk av fasttelefoner med tilsvarende kort.

Proxy CSCF

Når en brukeragent skal kontakte IMS systemet må den først sette opp en forbindelse med en P-CSCF. Når forbindelsen mellom brukeragenten og P-CSCF er satt opp, vil denne vedvare under hele forbindelsen. P-CSCF vil fungere som en mellomliggende server under hele forbindelsen, og all signalering fra brukeragenten går via P-CSCF.

I-CSCF

I-CSCF er kontaktpunktet innfor nettverket til en bestemt nettverksoperatør. Alle forbindelser som settes opp mot brukere innenfor dette nettverket må gå via I-CSCF

S-CSCF

S-CSCF er hovedserveren i forbindelse med IMS. S-CSCF har en rekke ansvarsområder, noen av disse er:

- Prosessere SIP Register meldinger.
- Autentisering av brukere.
- Ruting av trafikk til riktig P-CSCF, I-CSCF applikasjonsserver eller gateway.
- Ha støtte for å benytte ENUM til oversettelse mellom E.164 og SIP-URI.

Home Subscriber Server

I IMS skilles det mellom en privat og en offentlig brukeridentitet. Den private brukeridentiteten benyttes til registrering, autentisering, autorisasjon og administrasjon av brukeren ovenfor hjemmenettverket. Den private brukeridentiteten er en Network Access Identifier(NAI). I tillegg benyttes en eller flere offentlige brukeridentitet(er). Dette er brukeren sin kontaktinformasjon og kan enten være en SIP-URI eller et E.164 nummer representert som Tel-URI. Både den private og den offentlige identiteten må lagres i

ISIM. De to brukeidentitene refereres til som IM Private Identity(IMPI) og IM Public Identity(IMPU)

Hoveddatabasen for en gitt bruker kalles Home Subscriber Server(HSS).
Bindingen mellom den private og den offentlige brukeridentiteten lagers i HSS og kalles en serviceprofil. Denne profilen lastes ned av S-CSFC ved behov. HSS inneholder også funksjonalitet tilsvarende Home Location Register(HLR) og Autentication Center(AUC) slik at roaming med de tradisjonelle mobile tjenestene kan sikres. I IMS har HSS ansvaret for å holde oversikten over hvilke brukere som tilhører ulike S-CSCF.

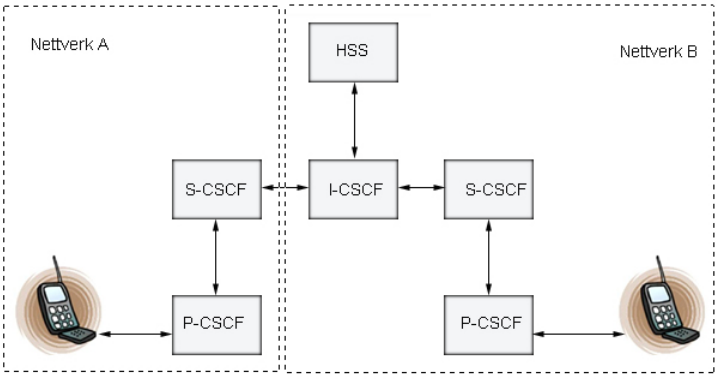
5.2 Samtrafikk

I dette kapittelet vil det bli gjennomgått en beskrivelse av samtrafikk i forbindelse med IMS. For å forenkle bildet noe er det tatt utgangspunkt i at registreringsprosessen for de ulike enhetene allerede er foretatt. All ruting innad i IMS skal baseres på SIP og SIP-URI benyttes som kontaktinformasjon. I IMS er det derfor nødvendig å ha en mekanisme med ENUM-funksjonalitet. I spesifikasjonen for IMS stilles det ingen krav til hvilken form for ENUM som skal benyttes. Det vil si at operatørene står fritt til å implementere sine egne databaseløsninger og ikke nødvendigvis benytte en åpen ENUM-løsning tilsvarende den beskrevet i kapittel 2.7.

Samtrafikk internt i IMS

Her kontakter brukeren i nettverk A brukeren i nettverk B med SIP-URI som kontaktinformasjon.

Figur 5.2. ”Samtrafikk mellom to IMS enheter” modifisert fra[6]

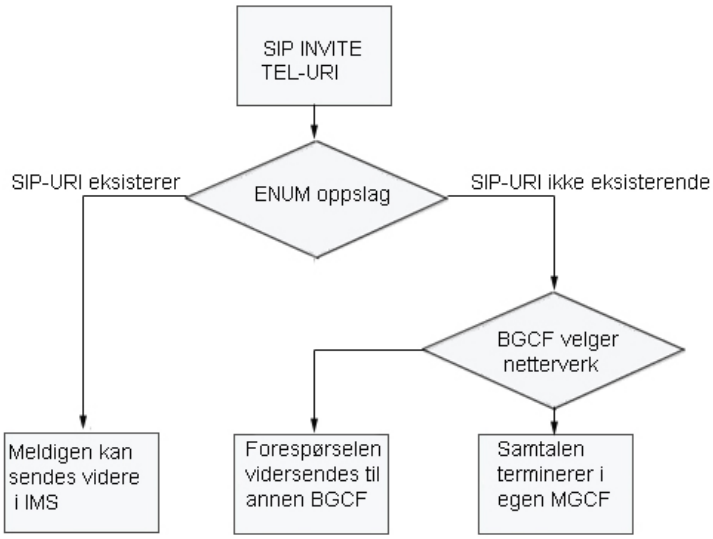


En sesjon startes ved at brukeren i nettverk A sender en INVITE melding til P-CSCF, deretter videresendes meldingen til S-CSCF. S-CSCF analyserer innholdet i meldingen og finner en passende I-CSCF server basert på adresseinformasjonen om bruker B. Når I-CSCF mottar meldingen kontaktes HSS for å finne hvilken S-CSCF som er ansvarlig for bruker B. S-CSCF prosesserer meldingen og videresender denne til P-CSCF. Tilslutt leveres meldingen til bruker B. Respons meldingen følger den samme stien tilbake, med unntak av at HSS trenger ikke å forespørres.

Samtrafikk fra IMS over i et linjesvitsjet nettverk.

Kommunikasjon innad i IMS foregår ved hjelp av SIP-URI. Hvis en S-CSCF isteden mottar en forespørsel med en TEL-URI(E.164) sjekker den derfor først om samtalen kan transporteres gjennom IMS-nettverket. Derfor gjøres det et oppslag i ENUM. Hvis det ikke finnes en SIP-URI for det aktuelle E.164 nummeret må samtalen rutes ut til en gateway. Et eksempel på dette er vist i figur 5.2.2.

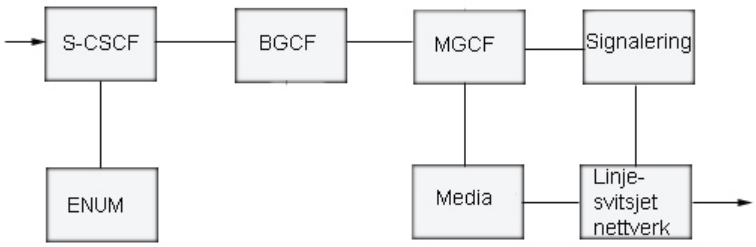
Figur 5.2.1 "Samtale fra IMS over til linjesvitsjet nettverk" modifisert fra[18]



Etter at S-CSCF har konstatert at samtalene må rutes ut til det linjesvitsjede nettverket kontaktes først en Breakout Gateway Control Function(BGCF). BGCF bestemmer hvilket nettverk som skal benyttes for å terminere samtalen ut i det linjesvitsjede nettverket.

Siden det er ønskelig å transportere samtalen lengst mulig innad i IMS samarbeider de ulike BGCF i ulike nettverk med hverandre ruting av samtaler. Den BGCG som blir valgt som termineringspunkt kontakter en Media Gateway Control Function(MGCF) i sitt lokale nettverk.

Figur 5.2.2 "Gateway funksjonalitet i IMS" modifisert fra[6]

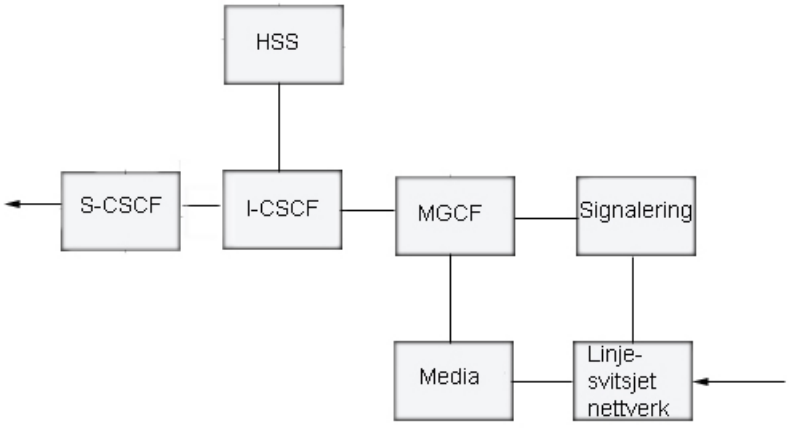


Virkemåten for MGCF er den samme som for en Media Gateway Controller som ble beskrevet i kapittel 2.5.

Samtale fra linjesvitsjet nettverk og inn i IMS

Her rutes samtalen i det linjesvitsjede nettverket helt inn til den aktuelle MGCF for den brukeren som blir oppringt.

Figur 5.2.3 "Samtale fra linjesvitsjet nettverk og inn i IMS" modifisert fra[6]



Signalering oversettes til SIP-signalering og sendes via MGCF til I-CSCF. I-CSCF benytter så HSS for å finne hvilken S-CSCF som er ansvarlig for den aktuelle brukeren.

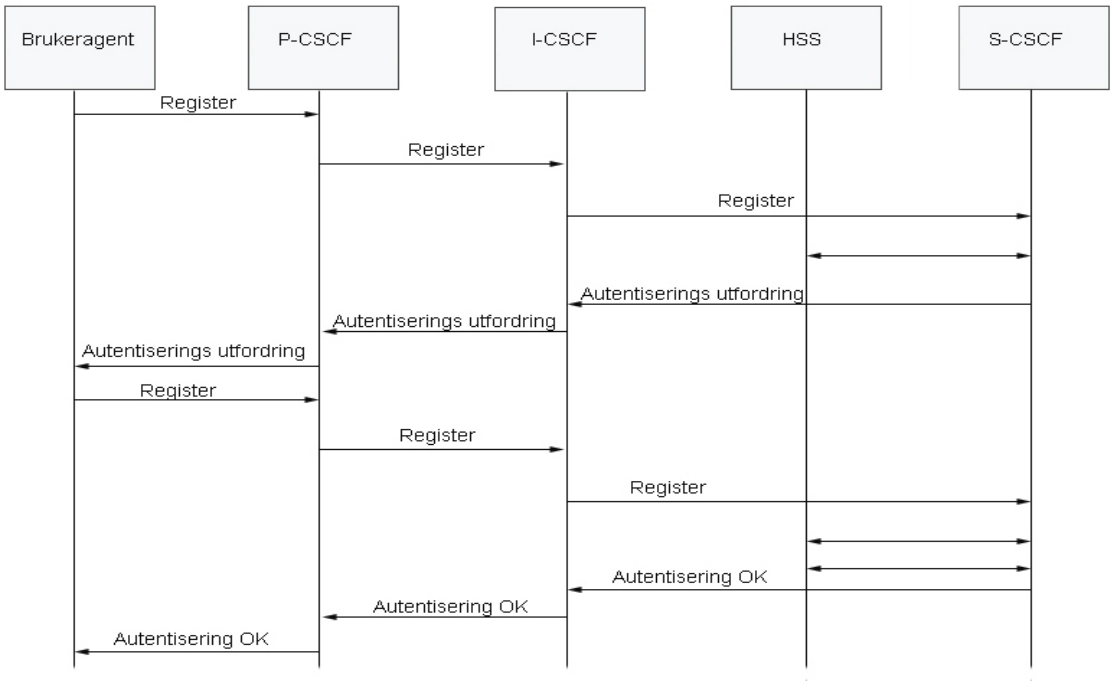
5.3 Sikkerhet

Sikkerheten i IMS er en basert på at ISIM-kortet og HSS har delte felles nøkler og et sett med tilhørende algoritmer. Prinsippene er omtrent de samme som benyttes for autentisering i UMTS.

Autentisering og nøkkelutveksling

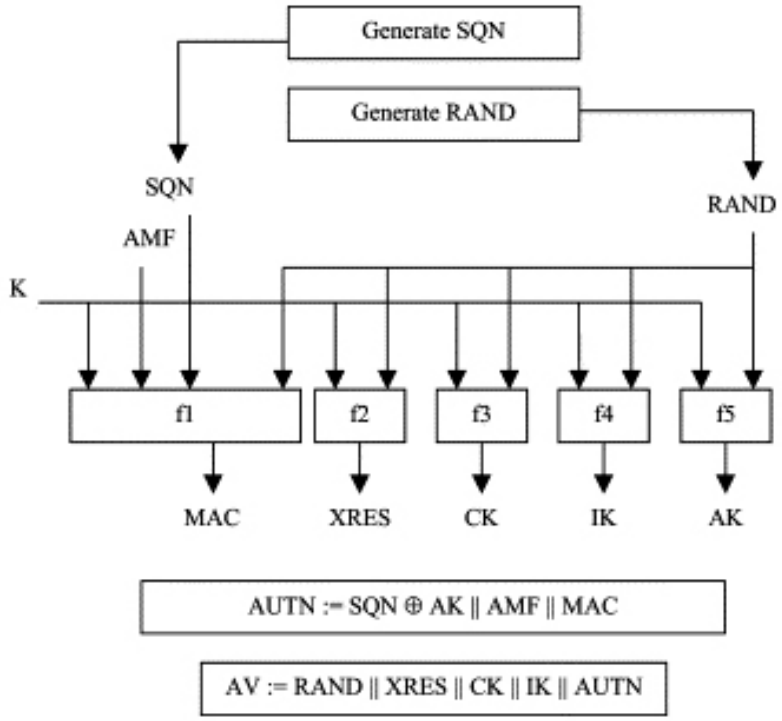
Før en bruker gis adgang til IMS er autentisering av brukeragenten nødvendig. Brukeragenten sender derfor en Register-forespørsel; *Register(IMPI, IMPU)*. Denne inneholder både den private og den offentlige identiteten til brukeren.

Figur 5.3 "Autentisering i IMS" modifisert fra [9]



I-CSCF mottar meldingen og tilegner deretter brukeragenten en bestemt S-CSCF og videresender Register-forespørselen hit. S-CSCF er deretter nødt til å kontakte HSS for autentiseringsinformasjon om brukereagenten. HSS sjekker at sammenhengen mellom IMPI og IMPU er gyldig. HSS sender deretter et sett med autentiseringsvektorer til S-CSCF som kan benyttes for autentisering av brukeragenten.

Figur 5.3.1 ”Generering av autentiseringsvektor i IMS” fra [47]



Fra figuren ser vi at det kreves fire ulike verdier for å danne en autentiseringsvektor. Dette er et sekvensnummer(SQN), en generert verdi(RAND), en delt nøkkel(K) og Authentication Management Field(AMF). AMF forteller hvilken nøkkel og algoritme som skal benyttes. Brukeragenten og HSS har som sagt på forhånd et sett med delte algoritmer og nøkler. Verdien XRES er en såkalt forventet respons. IK og CK er 128-bits nøkler for integritet og konfidensialitet. MAC, AK, AMF og SQN danner til sammen en autentiseringsstreng AUTN. Autentiseringsvektoren består av(RAND, XRES, IK, CK, AUTN). Hver vektor er gyldig for en enkelt autentisering.

En autentiseringstufordring sendes deretter tilbake til P-CSCF (IMPI, RAND, AUTN, IK, CK). Den forventende responsen (XRES) ligger lagret i S-CSCF og er ikke med i denne meldingen. P-CSFC lagrer nøklene(IK og CK) og sender forespørselen(IMPI, RAND, AUTN) videre til brukeragenten. Brukeragenten regner så ut sin egen MAC på samme måte som i figuren og sammenligner denne med den som ble tilsendt i AUTN.

Hvis både MAC og sekvenssnummer stemmer danner brukeragenten en responsmelding(IMPI, RES) som sendes tilbake til S-CSCF. I tillegg regner brukeragenten ut IK og CK. S-CSCF sammenligner deretter RES og XRES. Hvis disse stemmer overens er brukeragenten autentisert.

Konfidensialitet og integritet

IPsec i ESP transportmodus er den foretrukne metoden for å sikre konfidensialitet og integritet av SIP-trafikken i IMS. ESP er nærmere beskrevet i RFC2406[48]. For å realisere dette er det nødvendig å sette opp en sikkerhetsassosiasjon(SA). Her må det utveksles informasjon om hvilke algoritmer som skal benyttes for konfidensialitet og integritet og en sikkerhetsparameter indeks(SPI).
Vi så at det ble utvekslet nøkler i forbindelse med autentiseringsprosessen. Disse blir benyttet til å sørge for integritet i utvekslingen av disse initierende meldingene. Krypteringsnøklene dannes ved å benytte en ekspansjonsfunksjon på de nøklene som ble utvekslet i forbindelse med autentiseringen.

5.4 Nye tjenester i IMS

IMS introduserer noen tilleggskrav til tjenester i forbindelse med mobiltelefoni:

- Det skal være mulig å tilby tjenestekvalitet for en gitt sesjon.
- Det skal introduseres mer fleksible betalingstjenester slik at tilbydere skal kunne ta betalt for tjenester på et større grunnlag enn kun nedlastet datamengde.
- Det skal være større muligheter for å tilby mobilkunder nye tjenester basert på tredjepartsløsninger isteden for at dette skal styres kun av den aktuelle tilbyderen.

Eksempler på ”nye” tjenester er muligheter for fildeling, konferansesamtaler, chat-tjenester og ”push to talk”.

For å realisere dette benyttes applikasjonsservere eller mediaservere. Applikasjonsservere kan enten kan være i hjemmenettverket eller i nettverket til en tredjepart. Mediaservere er lokalisert i hjemmenettverket og benyttes i hovedsak til konferansetjenester.

6 Unlicensed Mobile Access

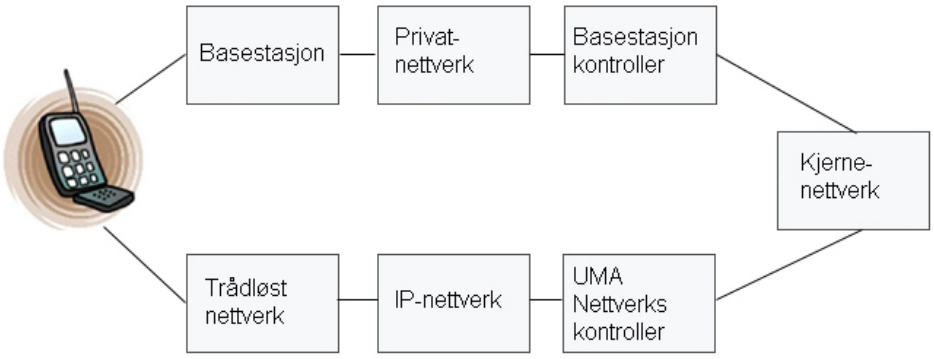
Unlicensed Mobile Access(UMA) er en teknologi som benyttes for å tilby kunder GSM og GPRS tjenester over trådløse datanettverk. Både Bluetooth og 802.11 standarden støttes. Hvis brukeren under en samtale beveger seg inn et område uten trådløst nettverk foretar mobilen en handover til det tradisjonelle mobilnettet uten at brukeren merker dette. Mobiltelefonen er altså avhengig av å ha støtte for begge teknologiene. Det er i det siste blitt lansert telefoner med UMA støtte fra blant annet Nokia.

Dette kapittelet vil gi en kort gjennomgang av UMA-teknologien basert på spesifikasjoner fra de deltakende aktører[2] sin samleside for UMA[16]. Fokuset vil være på hvilke nye løsninger som introduseres av UMA .

6.1 Arkitektur

Den overordnede virkemåten til UMA er ganske enkel. For å benytte GSM/GPRS tjenester kan brukeren enten benytte det tradisjonelle mobilnettet eller et trådløst nettverk.

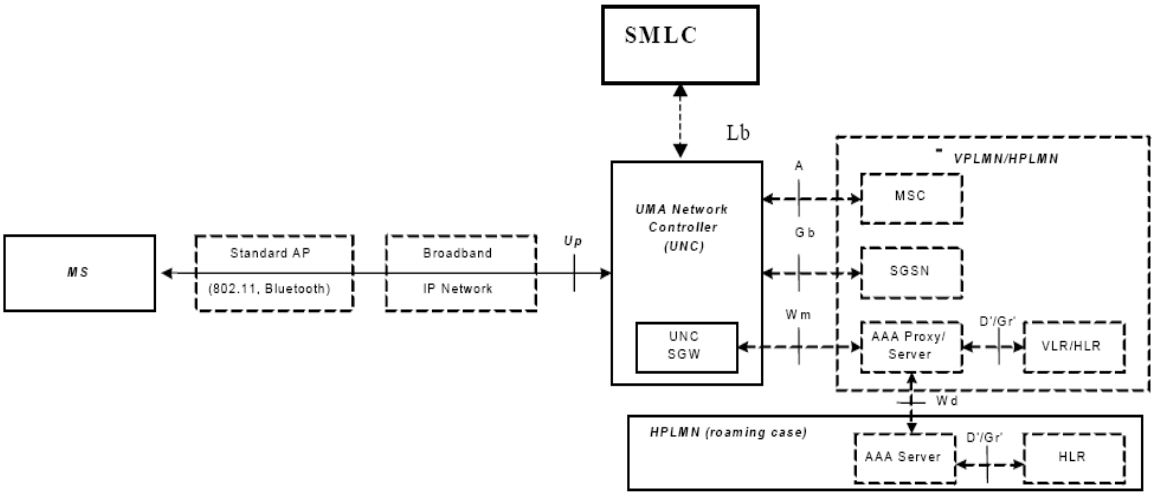
Figur 6.1 "Mobiltelefoni ved hjelp av UMA" modifisert fra[16]



Som vi ser fra figur 6.1 innebærer ikke et UMA-nettverk store forandringer i dagens teknologi. Den øverste stien er slik kjenner dagens mobilteknologi og vil derfor ikke bli noe nærmere forklart. Ved siden av utvidet funksjonalitet i selve mobiltelefonen er UMA- nettverkskontroller(UNC) det tillegget i nettverket som kreves av UMA.

[2] UMAer et samarbeidsprosjekt mellom følgende aktører:
Alcatel , British Telecom, Cingular, Ericsson, Kineto Wireless, Motorola, Nokia, Nortel Networks, O2 ,Research in MotionRogers Wireless,Siemens, Sony Ericsson og T-Mobile US

Figur 6.1.1 ”Enheter som kommuniserer med UMA-nettverkskontroller” fra[17]



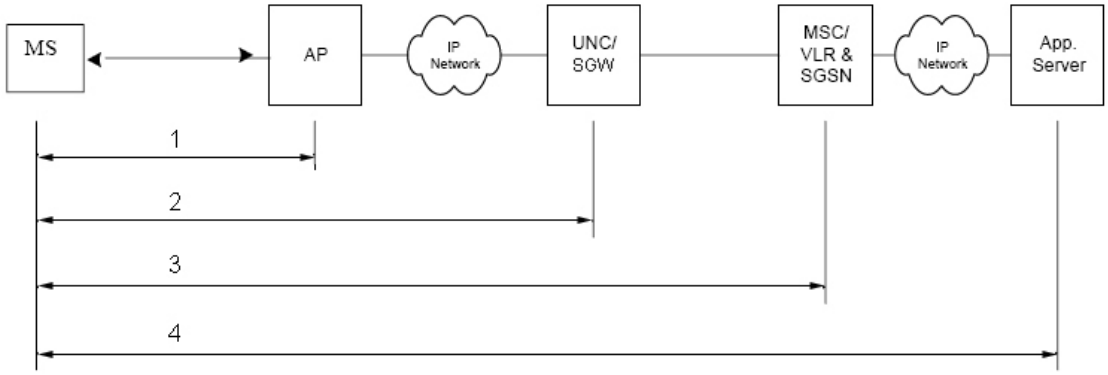
Av figuren ser vi at UNC kommuniserer ved hjelp av fem ulike grensesnitt.

- Up: Grensesnitt mot mobiltelefonen(MS). Når en forbindelse settes opp benyttes en Security Gateway(SGW) for autentisering og oppsett av en IP-Sec forbindelse ut til mobiltelefonen.
- Lb: Grensesnitt mot Serving Mobile Location Centre(SMLC). SMLC en lokaliseringstjeneste som skal benyttes i forbindelse med nødanrop.
- A: Grensesnitt mot Mobile Switching Center(MSC). Benyttes for linjesvitsjede telefonitjenester.
- Gb: Grensensitt mot Service GPRS Support Node(SGSN). SGSN har mye av samme funksjonalitet som MSC men SGSN benyttes mot pakkesvitsjede tjenester.
- Wm: Security Gateway benytter et Wm grensesnitt mot en lokal AAA(Authenticaction, Autorization, Accounting)-server i forbindelse med å kartlegge AAA-informasjon om brukeren. Serveren benytter Home Location Server(HLR) eller Visisted Location Register(VLR). En slik forespørsel om autentisering kan rutes over flere proxyservere avhengig av identiteten til brukeren.

6.2 Sikkerhet

UMA skal sørge for å sette opp en sikker forbindelse mellom UMA-nettverkskontrolleren og mobiltelefonen. Dette er vist som (2) i figur 6.2. De andre sikkerhetsforbindelsene(1, 3 og 4) som er vist i figuren er ikke tatt stilling til i UMA spesifikasjonene.

Figur 6.2 ” Sikkerhetsforbindelser i forbindelse med mobiltelefoni” modifisert fra [17]



1) Dette er forbindelsen mellom mobiltelefonen og aksesspunktet. Generelle metoder for autentisering og kryptering i slike tilfeller omtales nærmere i kapittel 7.3.
2) Denne delen må settes opp med en sikker forbindelse av UMA, og det er den eneste av de 4 forbindelse som UMA selv må sikre.
3) Denne forbindelsen beskriver den tradisjonelle autentiseringen og krypteringen av GSM eller GPRS. Dette er kun aktuelt når den øverste stien i figur 6.1 benyttes og er dermed ikke noe nytt i forbindelse med UMA.
4) Her beskrives eventuell tileggskryptering, for eksempel oppsett av en TLS forbindelse, hvis mobiltelefonen brukes til Internettjenester eller lignende. Dette er et generelt tema og ikke noe nytt som introudseres av UMA.

6.2.1 Autentisering og sesjonsoppsett

I forbindelse med autentisering av mobiltelefonen(MS) og UNC/SGW benyttes to ulike mekanismer; Extended Autehentication Protocol(EAP) og Internet Key Exchange (IKEv2)[49].

Extended Autehentication Protocol(EAP)

Virkemåten til Extended Authentication Protocol(EAP) er nærmere beskrevet i kapittel 7.3. I forbindelse med UMA kan to ulike former for EAP benyttes. Dette er EAP-SIM hvis MS benytter et SIM-kort ved GSM eller EAP-AKA hvis mobiltelefonen benytter et USIM kort som ved UTMS. Informasjonen i kortene benyttes av IKEv2 i forbindelse med autentisering.

Internet Key Exchange(IKEv2)

Internet Key Exchange(IKE) benyttes for autentisering og for å sette opp en sikkerhetsassosiasjon(SA) i forbindelse med IPsec. Internet Key Exchange versjon to er spesifisert i RFC4306[49] og er basert på elementer fra flere ulike standarder: Internet Key Exchange, Internet Security Association and Key Management Protcol(ISAKMP) og the Internet Domain of Interpretation(DOI). I tillegg inneholder IKEv2 nye elementer, blant annet utvidet støtte for NAT.

Etter at mobiltelefonen(MS) har etablert en forbindelse med aksesspunktet kontaktes UNC/SGW. UNC/SGW autentiserer seg med et X.509 sertifikat. Deretter foretas det en initierende nøkkelutveksling mellom MS og (UNC/SGW). UNC/SGW fungerer deretter som en mellomliggende enhet for en autentisering mot HSS/HLR. Dette er nærmere beskskrevet i kapittel 7.3..

6.2.2 Konfidensialitet og Integritet

IPsec i tunellmodus ved hjelp av ESP skal benyttes til å sikre all trafikk mellom MS og UNCI/SGW. Det foreslås to ulike sikkerhetsprofiler som brukeren kan velge imellom. Begge profilene må støttes av UNC/SGW mens MS må støtte minst en.

Tabell 6.2 "Ulike sikkerhetsprofiler i forbindelse med UMA" modifisert fra[17]

Profiler	Kryptering	Integritet	Tunnelmodus
Sikkerhetsprofil 1	3DES i CBC-modus, (RFC 2541)	HMAC-SHA1-96, (RFC 2404)	Ja
Sikkerhetsprofil 2	128 bits AES, (RFC 3602)	AES-XCBC-MAC-96, (RFC 3566)	Ja

Tabellen viser ulike algoritmer for kryptering og integritet i forbindelse med to sikkerhetsprofiler i forbindelse med UMA samt hvilke RFC'er som inneholder nærmere beskrivelser av disse algoritmene.

7 Utfordringer og krav til sikkerhet

Hvilke krav som skal stilles til sikkerhet vil være avhengig av hvilken type tjeneste en ønsker å realisere. Det vil også her være naturlig å skille mellom IP-telefoni som et supplement til fasttelefoni eller IP-telefoni som en erstatning for fasttelefoni. Det vil stilles strengere krav til tilbydere som leverer IP-telefoni som en erstatning for fasttelefoni. Dette gjelder blant annet oppetid, nødanrop, kommunikasjonskontroll og at kunden skal belastes riktig i forhold til bruk av tjenester. Tilbydere som ønsker å gi brukerne en offentlig telefonitjeneste må ha løsninger på disse områdende samt oppfylle øvrige krav som stilles i forbindelse med dagens PSTN-telefoni. Kapitel 7.1 inneholder en kort oppsummering av sikkerhet i forbindelse med SIP. En mer nøye gjennomgang av sikkerhetsutfordringer i forbindelse med denne type IP-telefoni er gjennomgått i den forrige oppgaven min[10].

I dette kapitelet vil det derfor bli fokusert på sikkerhetsproblematikk som kan oppstå ved en mer åpen tjeneste som ENUM. I tillegg vil det gjennomgått sikkerhet i forbindelse med trådløse nettverk og metoder for å sikre tjenestekvalitet.

7.1 Sikkerhet i forbindelse med SIP

I SIP-spesifikasjonen[11] omtales ulike former for sikkerhetsmekanismer i forbindelse med SIP. Tabell 7.1.1 inneholder en oppsummering av de ulike metodene og tabell 7.1.2 inneholder en kort oversikt over hvordan disse metodene kan benyttes i forbindelse med ulike trusler. En nærmere beskrivelse av dette finnes som sagt i[10].

Tabell 7.1.1 "Ulike sikkerhetsmekanismer i forbindelse med SIP

Sikkerhetsmekansime	Benyttes til	Signalering	Media	Nøkkeltuveksling
HTTP digest	Autentisering	*		Delte nøkler
S/MIME	Autentisering integritet og eventuelt konfidensialitet	*		PKI
TLS	Autentisering integritet og konfidensialitet.	*		PKI
IPsec	Autentisering integritet og konfidensialitet.	*	*	PKI.
Secure RTP	Autentisering integritet og konfidensialitet.		*	PKI

Tabell 7.1.2 ”Ulike typer angrep som kan rettes mot SIP”

Angrep	Konsekvenser	Konfidens.	Integritet.	Tilgjeng.	Beskyttelse
Avlytting av samtaler	Personlige eller bedriftsinformasjon kommer på avveie	*			IPsec SRTP
Samtalekontroll	Informasjon om samtalemønster og samtalepartere kommer på avveie.	*			S/MIME TLS IPsec
Falsk brukerregistrering	Tjenestenekting eller uautorisert bruk av tjenesten.	*	*	*	S/MIME TLS IPsec
Tjenestenektingsangrep	Tilgjengeligheten til tjenesten blir redusert eller totalt blokkert.			*	Benytte beskyttelse mot virus og ormer. Samt benytte moderne brannmurer

7.1.1 Utnytte svakheter i SIP INVITE

I kapittel 2.1.4 var det vist et eksempel på de ulike feltene i en SIP-INVITE melding. Ved å modifisere disse feltene er det mulig å fremprovosere *buffer overflow* angrep eller tjenestenektingsangrep mot ulike SIP-enheter. En test for å avsløre svakheter ved implementasjoner av SIP ble gjort av Oulo universitet i Finland[29]. Eksempler på endringer i SIP INVITE kan være:

- Overflow: Fylle opp med 128 Kbyte av ulike tegn som: /<>:@=
- Beskrive feil SIP versjon: Endre på SIP/2.0 som er standarden.
- Endre på SIP tag, tag brukes som identifikasjon av en sesjon.
- Endre på IP-adresseinformasjon
- Benytte feil beskrivelse av innholdet i meldingskroppen(application/sdp).

Etter at disse endringene benyttes på ulike måter i SIP-INVITE så sendes de modifiserte meldingene ut til ulike proxyservere. I testen beskrevet i [29] var det kun en av åtte leverandører som taklet alle de ulike modifikasjonene på en feilfri måte.

7.2 Spam over IP-telefoni

Spam over IP-telefoni(SPIT) blir regnet som en av de kommende problemene for IP-telefoni. I denne sammenhengen kan vi dele opp IP-telefoni i to hovedsegmenter. Den ene delen drives av offentlig godkjente tilbydere og er kommet som en arvtaker etter tradisjonell fasttelefoni. Den andre delen har stort sett vært basert på peer-to-peer løsninger over Internett og med liten mulighet for samtrafikk med det tradisjonelle telenettet. Etter hvert vil en kunne oppleve en større grad av sammensmelting mellom disse løsningene i forbindelse med rammeverk basert på ENUM.

Spam ble kjent gjennom e-post der det sendes ut e-post i stor skala til vilkårlige mottakere, ofte i forbindelse med reklame. Problemet ble så stort at det for enkelte ble svært tidkrevende eller nærmest umulig å sortere ut gyldig e-post blant all Spam posten. Det er altså et liknende problem en nå frykter at skal ramme IP-telefoni.

SPIT vil faktisk være et større problem en Spam, siden det vil være mye mer forstyrrende å bli stadig oppringt av falske telefoner enn å motta tilsvarende e-post. Dessuten vil SPIT-meldinger raskt fylle opp eventuelle telefonsvarere og minne med oversikt over ubesvarte anrop. Det er mye mindre jobb å slette meldinger fra en innboks for en e-post enn å sitte og lytte til en og en talemelding på en telefonsvarer. I tillegg vil SPIT samtaler ta opp kapasitet hos servere som normalt sett skulle behandlet gyldige samtaleoppsett. Resultatet kan i verste fall bli forsinkelser eller tjenestenekting for gyldige brukere.

Metoder for å forhindre Spam

I forbindelse med e-post har det etter hvert kommet ulike filtre for å sortere Spam og virus. Det enkleste av disse filtrene benytter en såkalt svarteliste(Blacklist) for å stoppe innkommende e-post ved å sammenligne forhåndsbestemte kriterier med headeren i e-posten. Ulempen ved å benytte denne teknikken er at listen hele tiden må oppdateres for å fange opp de nyeste formene for Spam. En slik svarteliste vil dessuten etter hvert bli veldig omfattende, og med det øker også faren for at gyldig e-post blir blokkert. En annen innfallsvinkel på problemet har vært å bygge opp et register av gyldige brukere(Whitelist). Dette innebærer at hvis en e-post ikke skal bli blokkert er avsenderen

avhengig av å være en del av mottakeren sitt register av gyldige brukere. Denne løsningen skalerer imidlertid dårlig i større skala, og dessuten forhindres mottaker å få e-post fra gyldige brukere som ikke er på listen. En løsning på dette har vært å la blokkert e-post ofte havne i en egen mappe, slik at brukeren har selv har mulighet til å sjekke om noe av den blokkerte e-posten er gyldig.

En annen mer dynamisk løsning er filtre som benytter en algoritme basert på Bayes Teorem[3]. Disse filtrene kombinerer begge metodene nevnt ovenfor, men avgjørelsen om en e-post er Spam eller ikke bestemmes ikke lenger ut fra enkle regler. Disse filtrene skanner både meldingshodet og innholdet i meldingskroppen etter såkalte Tokens som kan være bestemte ord, tegn, IP-adresser eller domenenavn. Hvert enkelt Token sjekkes deretter oppimot en database som inneholder oversikt over sannsynligheten for at de ulike Tokene opptrer i gyldig e-post eller Spam. Etter hvert som brukeren av e-post programmet avviser eller godkjenner ulik e-post vil sannsynlighetsfaktorene i denne databasen oppdateres og dermed også filteret sine egenskaper for å gjenkjenne Spam

Hvorfor er SPIT forskjellig fra Spam

Som nevnt begynner det nå å komme en del teknikker for å reduser Spam i forbindelse med e-post. Disse teknikkene er derimot ikke direkte overførbare til IP-telefoni på grunn av kravene til sanntid. Det vil si at oppsettet og kontrollen av en samtale må foregå innenfor en meget kort tidsperiode i forbindelse med IP-telefoni sammenlignet med e-post. Et annet forskjell er at det ikke er mulig å sjekke innholdet av en IP-telefoni samtale før den faktisk et satt opp i motsetning til Bayes metode for e-post der innholdet i meldingskroppen skannes.

[3] Bayes Teorem er en metode for å finn ut sannsynligheten for at en hendelse A inntreffer gitt at hendelse B har inntruffet.

$$P(A|B)=\frac{P(B|A)(P|A)}{P(B)}$$

Metoder for å forebygge SPIT

Det finnes flere ulike metoder for å undersøke hvorvidt det er sannsynlig at en innkommende samtale er SPIT eller ikke. For å undersøke dette kan det benyttes en del regler for å avgjøre dette.

- Hvilket nettverk er det samtaler initieres fra, og hvilke sikkerhetsassosiasjoner er det vårt nettverk har med dette nettverket. For eksempel er det mindre sannsynlig at et SPIT samtaler blir satt opp fra PSTN-nettet enn fra IP-nettet.
- Hvor godt brukerne er autentisert og hvilken forbindelse benyttes for å transportere signaleringen.
- Benyttes informasjon om brukeridentitet, eller ringes det med ønske om å fremstå med skjult nummer.
- Hvilke prisavtaler er gjeldende for samtrafikk mellom to ulike nettverk. Jo billigere aksess til et nettverk sine telefonbrukere er jo større er sannsynligheten for at uønskede samtaler initieres.

En annen mulighet å undersøke ringemønsteret til ulike brukere. Ved å bygge opp en oversikt over bestemte ringemønstre kan en etter hvert gjenkjenne signaturen til SPIT-samtaler. Eksempler på slike kriterier kan være følgende:

- Antall innkommende samtaler fra samme adresse per tidsenhet.
- Andelen av samtaler fra en bestemt adresse som blir besvart.
- Antallet brukere som ringes opp fra en bestemt adresse.
- Samme lengde på samtaler fra en bestemt kilde

Ellers finnes det løsninger der telefonibrukere bygger seg opp nettverk med andre brukere for å avgjøre hvilke samtalepartnere som kan regnes for å stole på eller ikke. En slik modell kan være at brukere oppdaterer en sentral tredjepart om brukere som regnes som trygge brukere. En annen mulighet er at det bygges opp en lenke av tillitt mellom ulike brukere der bruker A stoler på bruker B og dermed også alle brukere som er godkjent av bruker B og så videre.

7.3 Sikkerhetsmekanismer i forbindelse med trådløse nettverk

Sikkerhet i forbindelse med trådløse nettverk vil etter hvert spille en stadig større rolle ettersom en stadig større grad IP-telefoni samtaler foregår over denne typen nettverk. I dette kapitelet vil det derfor bli gjennomgått noen sikkerhetsaspekter ved trådløse nettverk.

Tabell 7.3.1 "Grad av sikkerhet i forbindelse med ulike sikkerhetsmekanismer"

Sikkerhetsmekanisme	Grad av sikkerhet
WEP	Liten
WPA	Begrenset
WPA2(802.11i)	Tilfredsstillende

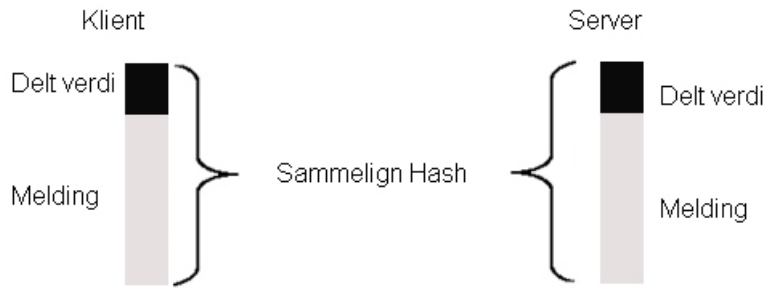
7.3.1 Wired Equivalent Privacy(WEP)

WEP er den originale standarden som ble beskrevet i forbindelse med autentisering og kryptering i forbindelse med 802.11.

Autentisering

Autentiseringen i WEP forgår ved at klient og server har en felles delt felles verdi. Serveren sender så en tilfeldig generert melding til klienten. Deretter beregner klienten en hashverdi på bakgrunn av denne meldingen og den delte verdien. Resultatet sendes tilbake til server som foretar den samme operasjonen og sammenligner verdiene. Hvis de er like er klienten autentisert.

Figur 7.3 "Bruk av delt verdi for autentisering"



Siden både meldingen og hashverdien sendes i klartekst og kan snappes opp av en angriper kan en ved hjelp av ”brute force” avsløre den delte verdien. WEP har heller ingen egne metoder for å distribuere den delte verdien og løsningen er derfor ofte å statisk implementere verdien i hver enkelt enhet.

Kryptering

Krypteringsalgoritmen som benyttes i WEP er RC4. Nøkkelengden er fra IETF spesifisert til 40 bit, men en del løsninger benytter en nøkkellengde på 104 bit. Dette kombineres med en initieringsvektor på 24 bit. Derfor refereres det ofte noe feilaktig til 64 eller 128 bits nøkkellengde. Siden det ikke er noe system for nøkkeldistribusjon gjøres dette ofte manuelt. Alle enhetene innenfor et aksesspunkt må benytte samme nøkler, og det støttes makismalt fire ulike nøkler. WEP regnes i dag som en usikker protokoll siden krypteringen enkelt kan knekkes med tilgjengelig programvare.

7.3.2 Wi-Fi Protected Access(WPA)

WPA er en videreutvikling av WEP og kommer med en del nye tillegg. Dette er blant annet Temporal Integrity Key Protocol(TKIP). TKIP benytter metoder for integritetssjekk av meldingene og en mer effektiv måte oppdatere nøkler på og gir dermed WPA noe bedre sikkerhet enn WEP. WPA i sin grunnleggende form kombinert med TKIP regnes allikevel ikke som en sikker mekanisme. Ved å kombinere WPA med Extended Authentication Protocol(EAP) vil en derimot få et mye sikrere rammeverk. Dette er nærmere beskrevet i kapittel 7.3.3.

7.3.3 Wi-Fi Protected Access 2(WPA2)/802.11i

WPA2 bygger på 802.11i standarden. Det nye i WPA2 er at 128-bits AES skal benyttes til kryptering isteden for 128-bits RC-4 som i WPA. TKIP er ikke godkjent i WPA2, og EAP skal derfor benyttes for autentisering.

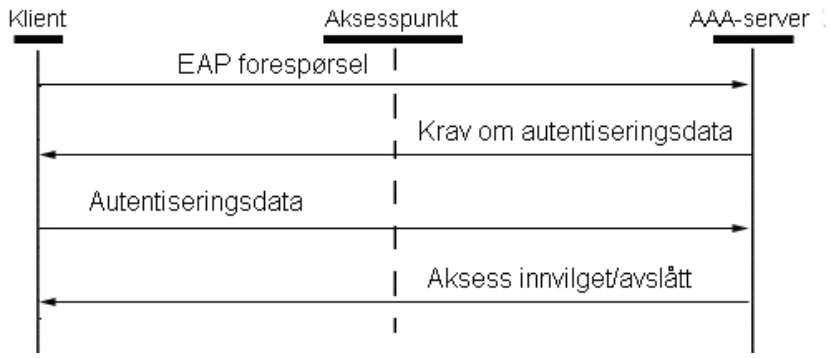
Autentiseringsmekanismer

I forbindelse med trådløse nettverk i IP-telefoni spiller autentisering en sentral rolle. Det er viktig at hver enkelt bruker av nettverket er i stand til å autentisere seg slik at uautoriserte brukere ikke er i stand til å initiere samtaler. Tradisjonelt sett har Peer-to-Peer(PPP) protokollen vært mye brukt i forbindelse med å sette opp en forbindelse mellom to noder i et nettverk. Autentiseringen i PPP har vært basert på et system der brukernavn og passord benyttes for å autentisere brukerne. Etter hvert ble det behov for en sterkere form for autentisering. Derfor ble det utviklet et tillegg til PPP standarden. Dette tillegget ble kalt Extended Authentication Protocol(EAP). EAP tilbyr en rekke ulike metoder autentisering, noen av disse vil bli nærmere gjennomgått senere.

Portbasert nettverksaksess (802.1x)

802.1x er en standard som ble designet for tradisjonelt Ethernet men den har etter hvert blitt tatt i forbindelse med trådløse nettverk. 802.1x benyttes for blant annet å transportere EAP-informasjon over et LAN. Ved å kombinere ulike sikkerhetsmekanismer kan 802.1x tilby både autentisering og kryptering.

Figur 7.3.1 "Klientautentisering ved hjelp av EAP"

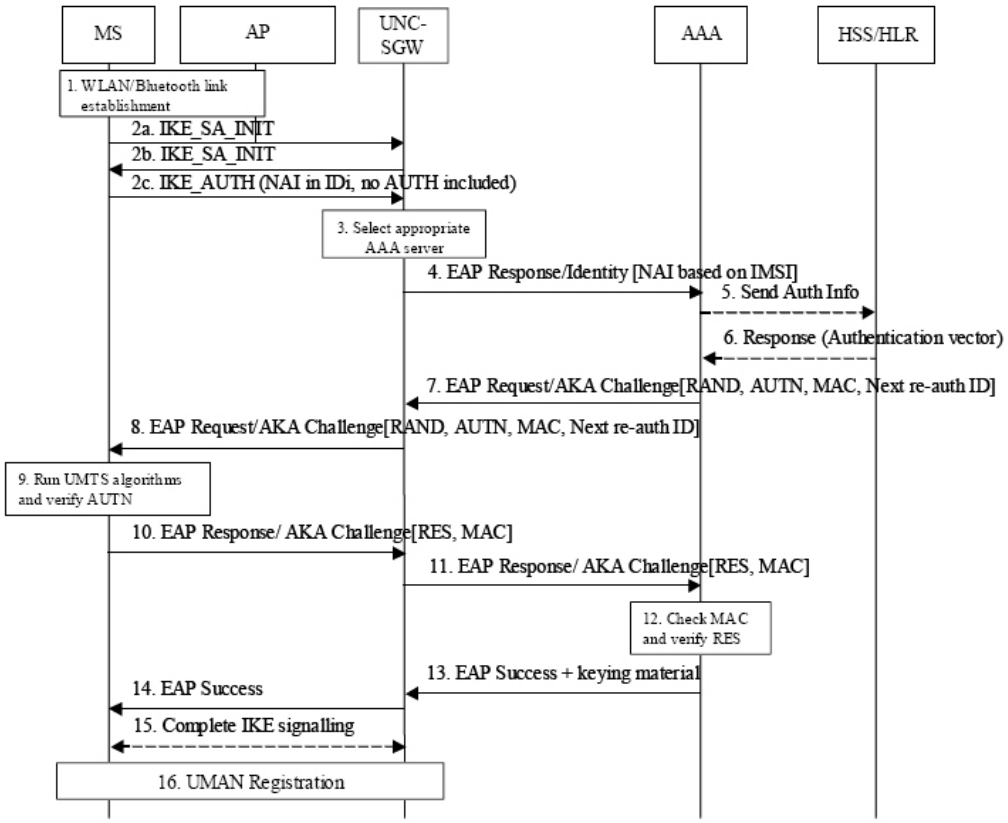


Figur 7.3.1 viser et forenklet oppsett av autentisering i foribindelse med EAP. Aksesspunktet kontakter alltid en AAA(Authenticaction, Autorization, Accounting)-server når aksess skal innvilges for klienten. AAA-servere benyttes i tillegg til formål som brukerpolicy, overvåkning av brukerprofiler og betalingstjenester. Den mest benyttede protokollen i forbindelse med AAA-servere er RADIUS, denne vil etter hvert bli erstattet av DIAMTER.

Autentisering med EAP i forbindelse med Unlicensed Mobile Acces

Det finnes mange ulike EAP autentiseringsmetoder. To av de mest vanlige er EAP-SIM og EAP-AKA. EAP-AKA benyttes i forbindelse med UMTS og autentiseringen som ble beskrevet i forbindelse med IMS bygger også på de samme prinsippene. I UMA benyttes enten EAP-SIM eller EAP-AKA, avhenigig av hvilke SIM-kort som benyttes. EAP-SIM og EAP-AKA baserer seg på forhåndslagrede nøkler i SIM-kortet.

Figur 7.3.2 "EAP-AKA i forbindelse med UMA" fra [17]



Fra figuren ser vi at Internet Key Exchange(IKE) benyttes for å sette opp en sikkerhetsassosiasjon(SA) i forbindelsen mellom den mobile enheten og en sikkerhetsgateway. Deretter foretas det en autentisering av den mobile enheten ved hjelp av en AAA-server som benytter informasjonen lagret i HSS/HLR. Prinsippene med bruk av en autentiseringsvektor er som sagt omtrent de samme som de som ble beskrevet i forbindelse med IMS.

Andre former for EAP

En etter hvert vanlig form for EAP autentisering er at det opprettes en ytre og en indre autentiseringstype. Først settes det opp en ytre TLS-forbindelse. Etter at TLS forbindelsen er satt opp kan en rekke andre autentiseringsmekanismer benyttes.

Tabell 7.3.2 ”Varianter av Extended Authentication Protocol(EAP)” modifisert fra [9]

Implementasjon	Server autentisering	Klient autentisering	Tunnel
EAP-TLS	Sertifikat	Sertifikat	TLS
EAP-FAST	Delt hemmelig verdi	Delt hemmelig verdi	
LEAP	Delt hemmelig verdi	Delt hemmelig verdi	
EAP-TTLS	Sertifikat	Sertifikat, PAP, CHAP, MS-CHAP-V2, EAP	TLS
EAP-PEAP	Sertifikat	Sertifikat, smartkort, MS-CHAP-V	TLS

- EAP-TLS: Her benyttes sertifikater for gjensidig autentisering av klient og server. Dette regnes som den sikreste formen av EAP. For å benytte EAP-TLS kreves det en infrastruktur for nøkkeldistribusjon(PKI) noe som gjør at enklere varianter av EAP heller benyttes.
- EAP-FAST og LEAP: Her autentiseres både klient og server ved hjelp av delte verdier. En mulig fare ved denne varianten er at en angriper kan snappe opp hash-verdien av passordet. Deretter kan det brukes ulike ”brute-force” verktøy for å avsløre passordet.

- EAP-TTLS: I motsetning til EAP-TLS krever ikke EAP-TTLS at klienten autentiserer seg med et sertifikat, men det er støtte for dette hvis det er ønskelig. EAP-TTLS kan benyttes sammen med ulike former for indre autentisering.
    - Password Authenticaction Protocol(PAP): Dette er en enkel ID/Passord tostegs autentiseringsmekanisme. Denne regnes ikke lenger som sikker og har av IETF blitt oppgradert av Challenge Handshake Autentication Protocol(CHAP)
    - CHAP: Her er det tatt utgangspunkt i at klient og server har en delt hemmelig verdi. Autentiseringen starter med at serveren sender en verdi til klienten. Klienten benytter så en enveis hash-funksjon for å beregne en verdi på grunnlag av den delte hemmelige verdien og verdien den har mottatt fra serveren. Hash-verdien oversendes så serveren som utfører samme operasjon som klienten og sammenligner hash-verdiene. Hvis de er like er klienten autentisert. Microsoft har også en variant av denne mekanismen som kalles MS-CHAP. Denne har ethvert blitt utvidet med en MS-CHAP versjon 2
- EAP-PEAP: Her autentisers klienten serveren med et sertifikat mens klienten kan benytte ID/Passord eller eventuelt løsninger med smartkort.

7.4 Tjenestekvalitet

I forbindelse med telefoni er det en del krav som stilles en telefonitjeneste fra et brukersynspunkt. Et av de viktigste kravene er at tjenesten er tilgjengelighet. Dette er avhengig av oppetiden til tjenesten. Andre krav til en tjeneste er lav forsinkelse ved oppsett av samtalen og kvaliteten på selv talen. Tre ulike faktorer er med på å påvirke tjenestekvaliteten:

1) Forsinkelse: Tiden en pakke bruker for å transporteres fra brukeragent A til brukeragent B. Akseptabel enveis forsinkelse i forbindelse med IP-telefoni er 150 ms.
2) Jitter: Variasjoner i forsinkelsen mellom hver enkelt pakke. Dette gjør at samtalen oppfattes som hakkete.
3) Pakketap: Antall pakker som ikke når frem til mottakeren.

Metoder for å kontrollere tjenestekvaliteten over IP-nettet kan en deles inn i Integrated Services(IntServ) og Differentiated Services(DiffServ).

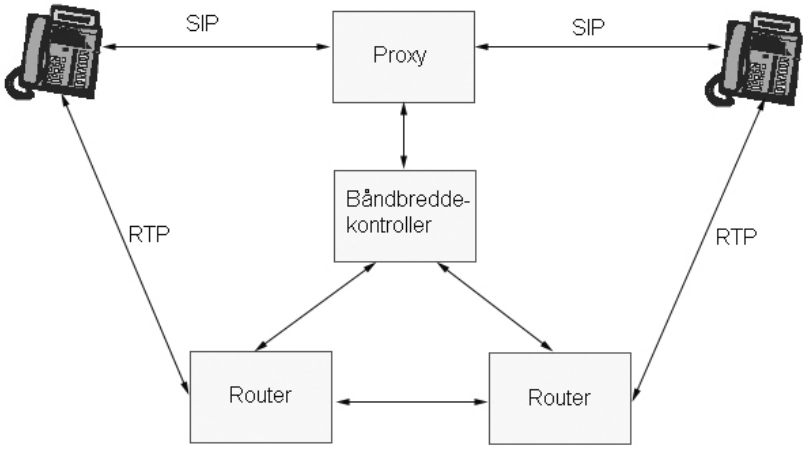
Integrated Services og Resource Resvervation Protocol

En mye brukt protokoll i forbindelse med IntServ er Resource Reservation Protocol(RSVP). RSVP fungerer ved at mellomliggende rutere må reservere en gitt båndbredde til deltakerne innenfor en enkelt sesjon. RSVP skaper mye ekstra trafikk i nettet siden nye reserveringer må gjøres for hver sesjon. Siden IP-telefoni samtaler ofte kan være korte eller aldri ble besvart vil RSVP i forbindelse med IP-telefoni være veldig ressurskrevende. Det finnes i dag enkelte løsninger som benytter RSVP i forbindelse med IP-telefoni, men på grunn av ressursbruk vil neppe RSVP bli benyttet i stor skala sammen med IP-telefoni. RSVP er nærmere beskrevet i blant annet RFC2210[51] og RFC4230[52].

Differentiated Services

DiffServ deler trafikken inn i ulike klasser og gir ulik prioritet til de ulike klassene, og DiffServ skalerer derfor bedre enn IntServ i forbindelse med IP-telefoni. For å realisere denne teknologien i forbindelse med SIP trenger proxyserveren å samarbeide med en enhet for båndbreddekontroll samt ruterne i nettverket. Løsningen er derfor mest nyttig når en tilbyder ønsker å tilby taletjenestene en bedre kvalitet enn for datatjenestene. Dette skyldes at rene datatjenester som regel ikke er så sensitive for forsinkelser som telefonitjenster.

Figur 7.4 "Bruk av Differentiated Services i forbindelse med SIP"

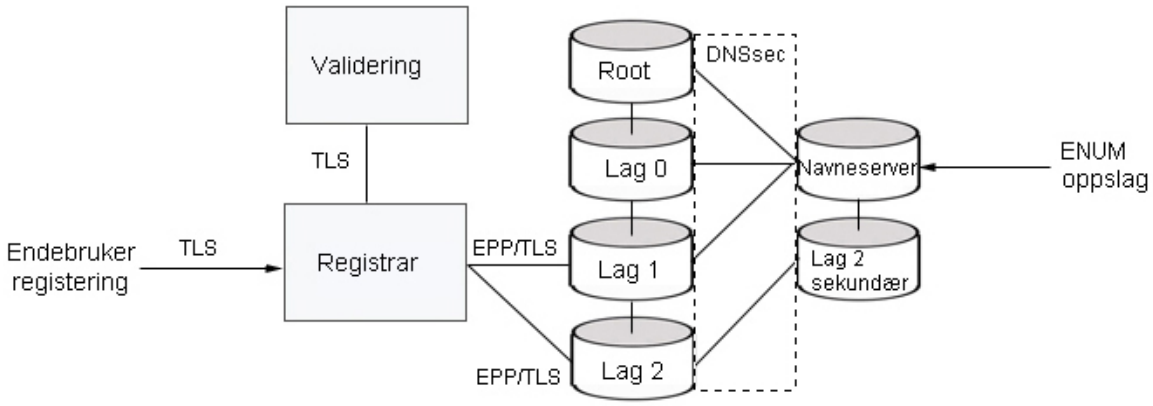


Når en SIP forespørsel ankommer proxyservern stilles det krav om hvilken båndbredde som er ønskelig fra brukeragenten. Proxyserveren videresender så denne informasjon sammen med informasjon om brukeragentene til en båndbreddekontroller. Bånbreddekontrolleren får da ansvaret for å sette opp mediakanalen og reservere båndbredde hos ruterne.

7.5 Sikkerhet i forbindelse med ENUM og DNS

I kapittelet om ENUM ble det gjennomgått et eksempel på hvordan er endebruker kan registere E.164 nummeret sitt ved hjelp av ENUM. I dette kapittelet vil det bli sett nærmere på hvilke sikkerhetsmekanismer som kan benyttes i forbindelse med en slik registrering og i forbindelse med ENUM og DNS generelt.

Figur 7.5 "Sikkerhetsmekansimer i forbindelse med ENUM" modifisert fra[25]



Figuren viser et forslag til hvilke sikkerhetsmekansimer som kan benyttes for å beskytte kommunikasjonen mellom sentral elementer i forbindelse med ENUM.

Validering av endbruker

Endebrukeren kan for eksempel kontakte en webserver hos en registrar når vedkommende skal registrere sitt E.164 nummer i ENUM. Denne forbindelsen bør være sikret i form av TLS slik at sensitiv informasjon ikke sendes ukryptert over Internett.
I neste fase er en nødt til å bekrefte at brukeren faktisk har rettigheter til å registrere det aktuelle E.164-nummeret i ENUM. Brukeren er med andre ord nødt til å ha en eller annen form for bevis ovenfor registrar. ETSI[24] foreslår flere mulige løsninger på dette.

- Tjenestetilbyderen benytter en tredjepart for å validering av endebrukeren. Dette kan for eksempel gjøres ved at tilbyderen av en ENUM-tjeneste kontakter PSTN-tilbyderen som har ansvaret for E.164 nummeret og får en bekreftelse på at endebruker er den rettmessige eier av E.164 nummert. Et annet forlslag er at

endebrukeren får utstedt et digitalt sertifikat som bekrefter bindingen mellom E.164 nummer og identitet.

- Et annen løsning som forslås er at brukeren må sende inn en for eksempel en telefonregning eller annet form for bevis til ENUM-tilbyderen for å bekrefte eierskap ovenfor E.164 nummeret.

Legg inn informasjon i DNS

En registrar må kontakte et register for å legge inn ny informasjon i DNS eller for å gjøre endringer på tideligere innlagt informasjon. I figur 7.5 er dette vist som forbindelsene mellom registrar og lag1/lag2. Tideligere har det ikke vært noe bestemt rammeverk for hvordan denne informasjonen skal overføres. Derfor ble Extensible Provisioning Protocol(EPP) i RFC3730[26] utarbeidet av IETF som en løsning på dette. EPP er en XML basert protokoll som behandler informasjonen som objekter. Denne informasjonen bør sendes over en sikker forbindelse. I RFC3734[27] spesifiseres det hvordan EPP kan benyttes over en TLS forbindelse.

Oppslag i DNS

Ved å basere en telefonitjeneste på å lagre E.164 informasjon i DNS som med ENUM vil også eventuelle trusler som og angrep rettet mot DNS også indirekte blir angrep mot selve telefonitjenesten. Det er veldig mange forskjellige trusler og ulike typer angrep som kan benyttes i forbindelse med DNS. Eksempel på dette kan være:

- **Endring av pakker:** DNS-pakker som sendes ubeskyttet over UDP uten noen form for signering eller kryptering er en potensiell fare. En angriper kan endre innholdet i meldingen på veien mellom mottaker og avsender. Hvis det er gjort et oppslag for å finne adresseinformasjon om et E.164-nummer kan altså svaret tilbake være falsk. Dette første som skjer er at brukeren som skal ringes opp ikke blir tilgjengelig. Deretter har også angriperen mulighet til å sende samtaler videre til den destinasjonen angriperen selv ønsker.

- **Endring av pakker kombinert med ”brute force”**: For å endre på pakker slik som i eksemplet vist ovenfor er angriperen nødt til å være i stand til å ha tilgang til selve pakkestrømmen for å identifisere portnummeret og DNS-identiteten til offeret. Hvis angriperen ikke har tilstrekkelig tilgang til mellomliggende nettverket vil ikke et slikt angrep være mulig. En annen mulighet er derfor at angriperen selv gjetter seg til denne informasjonen. Siden DNS benytter et statisk portnummer på serversiden trenger en angriper bare å finne kombinasjonen mellom DNS-identiteten og portnummet på klientsiden. Siden begge disse er på 16 bit gir dette 2^32 mulige kombinasjoner. Falske DNS-pakker kombinert med ”brute force” er derfor en mulig trussel.

- **Forgiftning av DNS cache:** En Internettkunde vil ofte bruke en DNS-server som er satt opp av tjenestetilbyderen. For å slippe å gjøre de samme DNS-oppslagene om og om igjen lagres disse i serveren sin cache. Forgiftning av DNS cache går ut på at en angriper sender falske meldinger til en DNS-server. Hvis serveren ikke har gode metoder for å sjekke om disse meldingene kommer fra en autorativ server vil falsk informasjon bli lagret i serveren sin cache. Dermed vil alle brukerne av denne serveren også motta falsk informasjon.

Dette var bare noen eksempler på ulike angrep. DNS Security Extensions(DNSsec)sine hjemmesider[15] inneholder en rekke ulike eksempler på angrep mot DNS.

The Domain Name System Security Extensions (DNSSEC)

Formålet med DNS Security Extensions (DNSsec) er å forhindre ulike typer angrep mot DNS, deriblant angrep mot DNS cache som beskrevet ovenfor. Dette realiseres ved å benytte digitale signaturer for autentisering og integritet i forbindelse med DNS-meldingene. DNSsec benytter en modell med offentlig og private nøkler. En autorativ DNS-server signerer de ulike sonene med sin private nøkkel. Serveren som mottar meldigene benytter deretter den offentlige nøkkelen til den autorative serveren for å verifsere at signaturen er gyldig.

Problemstillinger satt opp av ETSI

European Telecommunications Standards Institute(ETSI) tar i [24] opp noen ulike sikkerhetsutfordringer de ser for seg i forbindelse med ENUM. Truslene som nevnes her er blant annet:

- Registrering av falsk brukerinformasjon: Dette innebærer at noen legger inn falsk informasjon på vegne av andre brukere. Det vil da foreligge feil informasjon i DNS for det akutelle E.164-nummeret. En kan da få situasjoner der en bruker ringer ut til et E.164 nummer og samtalen blir satt opp, men endebrukeren i den andre enden er en annen enn eieren av nummeret. Det kan også oppstå situasjoner der en blir oppringt av et kjent nummer, men brukeren i den andre enden er en helt annen en eieren av nummeret.
- Registrering av E.164 nummer i ENUM uten brukers tillatelse: Her vil innkommende samtaler inn til en bruker potensielt kunne bli rutet via nettverk eller opp mot tjenester som brukerne selv ikke ønsker å benytte.
- Flere ulike registreringer for samme nummer: En ser for seg et problem med at det utvikles åpne andre systemer for nummeroppslag uten den samme strenge kontrollerte hierarkiske modellen som benyttes i ENUM. En vil da potensielt kunne få inkonsistens i forbindelse med navneoppslag.
- Lokasjon av DNS-servere: Det kan stilles spørsmålstegn ved om det er noen god løsning å basere en telefontjeneste, i et område eller land, på at DNS servere i et helt annet område er tilgjengelige.

8 Diskusjon

I denne delen vil det bli en oppsummering og diskusjon rundt noen av de sentrale temaene som er omtalt tideligere i oppgaven.

Hittil i oppgaven har det blitt presentert en del ulike teknologier i forbindelse med overgangen fra PSTN til IP-telefoni. Dette er et område som er under stadig utvikling, men tendensen er veldig klar. En stadig større del av telefonisamtalene vil foregå over IP-baserte nett isteden for de tradisjonelle PSTN-nettene. Denne utviklingen forsterkes ytterligere ved introduksjonen av ENUM som sørger for en enda sterkere binding mellom telefonitjenester og IP-adresser.

E.164 nummeret kan altså benyttes til ruting uten å gå omveien om PSTN. Når dette kombineres med SIP kan en potensielt få enklere og mer åpne løsninger i forbindelse med ruting enn dagens løsninger som baserer seg på lukkede intelligente nettverk i forbindelse med PSTN.

Samtidig ser en at denne åpne utviklingen reverseres noe når telefonioperatørene nå introduseres neste generasjons nettverk, som nå er i starfasen i form av IMS. Her ønsker tilbyderne å ta tilbake noe av kontrollen ved å flytte mer av funksjonaliteten og intelligensen tilbake igjen til kjernenettverket.

8.1 Samtrafikk

Hvis en IP-telefonisamtale skal rutes over et PSTN-nett er samtalen nødt til å passere gjennom minimum to media gatewayer, en for å komme ut på PSTN-nettet og en for å komme tilbake på IP-nettet igjen. Hvis brukerne i tillegg benytter tradisjonelle PSTN-telefoner trengs det ytterligere to gatewayer for å sørge for en analog til digital konvertering hos endebrukeren. I oppgaven har det blitt gjennomgått noen ulike former for gateway funksjonalitet og rammeverk som kan benyttes for å sikre en slik samtrafikk. Selv om en ser at stadig flere brukere går over til IP-telefoni vil det fortsatt være et behov for samtrafikk med PSTN-nettet i lang tid fremover. Mye ressurser innen IP-telefoni benyttes derfor på å gjøre IP-telefoni best mulig kompatibelt med PSTN.

Generelt er bruk av PSTN-nettet i forbindelse med IP-telefoni lite fordelaktig med tanke på tjenestekvalitet og samtalekostnader. I forbindelse med tjenestekvalitet kan en generelt kan en si at en samtale bør rutes gjennom færrest mulig gatewayer. Dette skyldes at hver gateway er nødt til å prosessere meldingene. Jo flere gatewayer en IP-pakke er nødt til å passere jo større er sjansen for redusert tjenestekvalitet i form av forsinkelse eller jitter. Dessuten påløper det samtalekostnader i forbindelse med samtaleavgifter ved bruk av PSTN-nettet.

Ved introduksjon av ENUM vil allikevel mye av markedet potensielt kunne forandres. En mye større andel av IP-telefonisamtaler vil kunne rutes uavhengig av gatewayer og PSTN-nettet. En kan derfor se for seg en utvikling med et marked der flere tilbydere av IP-telefoni samarbeider om å legge ut informasjon om sine endebrukere ved hjelp av ENUM, slik at de slipper unna PSTN-ruting kostnader. Dette vil bety et betydelig inntektsttap for de tilbyderne som i dag har kontrollen over PSTN-nettet. Tilbydere som har kontroll over PSTN-nettet vil naturligvis ikke ønske en slik utvikling, og ønsker derfor å forsinke prosessen.

8.2 ENUM basert på e.164.arpa (Public ENUM)

Dette er den løsning som har blitt gjennomgått tideligere i denne oppgaven. Kravene er at e164.arpa skal benyttes til å lagre E.164 numre i DNS og at det administrative skal kontrolleres av de aktuelle myndigheter i de ulike landene. En slik løsning omtales ofte som *Public ENUM* eller *User ENUM*. Post- og teletilsynet omtaler denne løsningen som Sluttbruker ENUM. I forbindelse med selve registreringen er det enighet om en løsning som kalles ”opt-in”, det vil si at brukerne selv skal kunne bestemme om deres nummer skal bli tilgjengelig via ENUM. Registreringen kan som tideligere vist foretas av selve brukeren eller en tiltrodd tredjepart i form av for eksempel en tjenestetilbyder. Denne teknologien kan i utgangspunktet benyttes av alle som ønsker å benytte sitt E.164 nummer som kontaktinformasjon også på Internett. I utgangspunktet kan dermed hver enkelt bruker sette opp sin egen IP-telefonitjeneste. Isteden for at hver enkelt bruker setter opp sin egen tjeneste er det mer sannsynlig at dette blir gjort av en tiltrodd tredjepart. Eksempler på dette kan være universiteter eller bedrifter som lanserer en IP-telefoni tjeneste på vegne av sine medlemmer/ansatte. Hvis de har den nødvendige infrastrukturen på plass vil de kunne etablere sin egen telefonitjeneste uavhengig av ulike tilbydere.

En slik løsning er altså basert på at hver enkelt bruker er registrert i ENUM. Det vil derfor bare være mulighet for samtrafikk med en bestemt gruppe brukere. Det er verdt å merke seg at en slik løsning ikke kan defineres som en offentlig telefonitjeneste. En slik tjeneste blir derfor ingen erstatning for dagens telefonitjeneste, men må heller betraktes som et supplement. Så lenge deltakerne drifter sine egne nettverk vil det ikke være direkte samtaleutgifter i forbindelse med en slik løsning. Av mer kommersielle løsninger kan en se for seg mer rene Internett løsninger hvor aktører tilbyr IP-telefoni til sine kunder hvis de tillater nummeret sitt registrert i ENUM.

En kan også se for seg hybride løsninger der dagens tilbydere tilbyr ENUM som en tilleggstjeneste til sine kunder. Kunden har da et abonnement på en vanlig telefonitjeneste, men kan i tillegg kontaktes via ENUM-tjenesten. Hvis en IP-telefonitilbyder velger å tilby ENUM for registering av E.164 numre for sine kunder vil

det være lite sannsynlig at brukerne i stor skala selv kontakter tilbyderen og ytrer et ønske om at de vil ha nummeret sitt registrert i ENUM. Det er derfor trolig at initiativet må komme fra tilbyderen, men dette forutsetter en godkjenning fra kunden. Denne løsningen er altså basert på at endebrukeren selv avgjør om informasjonen om brukeren sitt E.164 nummer skal offentliggjøres i DNS. Mange tilbydere ønsker derfor en mer tilbyderkontrollert løsning. Et eksempel på en slik løsning er Infrastruktur ENUM.

8.3 Infrastruktur ENUM

Her er kontrollen av nummeret overlatt fra brukeren til tjenestetilbyderen. Det er altså den tilbyderen som på et gitt tidspunkt har ansvaret for en abonnent som har muligheten til å legge inn ENUM informasjon i DNS. Denne løsning fokuserer altså mest på hvordan ENUM kan brukes som et rutingverktøy for en offentlig telefonitjeneste. Dette er noe forskjellig fra Sluttbruker ENUM der fokuset er mer rettet på hvordan E.164 nummeret kan benyttes som kontaktinformasjon for Internettbasert telefoni.

Tabell 8.1 "Sluttbruker ENUM og Infrastruktur ENUM" modifisert fra[23]

Problemstilling	Sluttbruker ENUM	Infrastruktur ENUM
Hvem skal tilby informasjon	Brukere av tjenesten	Tjenestetilbydere
Hvem har tilgang til å legge inn informasjon?	Alle skal ha tilgang til å legge inn informasjon vedrørende sitt eget E.164 nummer. Dette kan enten gjøres av eieren av nummeret eller en tiltrodd tredjepart.	Tjenestetilbyderen legger inn informasjon på vegne av sine kunder. Kunden blir ikke forespurt.
Definert domene	e164.arpa	Ikke bestemt

Vi ser at ved infrastruktur ENUM vil endebrukerne ikke ha muligheten for å selv å bestemme hvorvidt informasjon skal legges ut. Dette valget skal tas av tjenestetilbyderen. Det vil ofte oppleves som tungvindt for de ulike tilbyderne at det er kunden som må avgjøre om nummeret skal benyttes i ENUM slik som i varianten med Sluttbruker-

ENUM. Det er ennå ikke fastsatt hvilken DNS-løsning som skal benyttes i forbindelse med Infrastruktur ENUM.

8.3.1 Privat ENUM

Sluttbruker ENUM og Infrastruktur ENUM er to varianter der adresseinformasjon gjøres globalt tilgjengelig via DNS. I tillegg til disse to er det en ENUM løsning som ofte omtales som Privat ENUM. Her benytter tilbyderne sine egne lukkede ENUM løsninger. Her ønsker tilbyderne å begrense informasjon om sine abonnenter. Tilbyderne ønsker å benytte prinsippene fra ENUM teknologien for å unngå ruting via PSTN. Samtidig ønsker de å hemmeligholde adresseinformasjon om sine egne kunder og sin egen nettverkstopologi.

8.4 IP-telefoni over trådløse datanettverk

I oppgaven ble det teknologien Unlicenesed Mobile Acces gjennomgått. Her så vi at det var mulig å kombinere telefoni over 802.11 eller Bluetooth med tradisjonell mobiltelefoni. En slik løsning er på mange måter en midlertidig løsning. Manglene ved denne løsningen er at hver gang brukeren skal forflytte mobiltelefonen mellom to 802.11 nettverk er det nødvendig å overføre en handover til det GSM-nettet for deretter å utføre en handover tilbake til det nye 802.11 nettverket. GSM-nettet må derfor involveres i alle former for handover.

Etter hvert som veksten av områder med trådløs dekning øker vil det bli et større marked for slike løsninger. Et stort marked for dette vil det antageligvis ikke bli før det er teknologi tilstede for å få en myk handover mellom de ulike 802.11 nettverkene uten at GSM-nettet trenger å involveres. UMA kan derfor best beskrives som en evolusjon snarere enn noen revolusjon.

8.5 Videre utvikling

Sånn situasjonen er i dag kan en dele IP-telefoni opp i to forskjellige retninger. Den ene er styrt av godkjente telefonitilbydere og er tilrettlagt for alle-til-alle kommunikasjon. Det vil si at de ulike tilbyderne er avhengig av ha samtrafikkmuligheter slik at endebrukeren kan benytte E.164 for samtrafikk både med PSTN og andre IP-telefonibrukere. Denne formen for IP-telefoni kan altså sees på som en erstatning for tradisjonell fasttelefoni. Ved å benytte Infrastruktur ENUM eller private former for ENUM vil de ulike tilbyderne være i stand til å levere en billigere tjeneste siden større deler av samtalene kan rutes uavhengig av PSTN. Også en mekanisme som TRIP vil føre til rimeligere samtrafikkløsninger siden tilbyderne i større grad har muligheten til å utveksle gateway informasjon. Bruk av TRIP vil kunne føre til at samtalen kan rutes i lengre i IP-nettet fordi det blir lettere for tilbyderne å finne gunstige PSTN-termineringspunkter.

På den andre siden har vi løsninger som Skype. Her blir Internettkunder tilbudt telefoni til gunstige priser. En Internettelfonitjeneste kan også utvides med videotelefoni, fildeling, chatting osv. Dette er altså ikke en fullverdig offentlig telefonitjeneste, men et supplement. Et viktig aspekt ved dette er allikevel at verdien av en offentlig telefonitjeneste blir redusert, siden kundene nå ikke lenger er avhengige av en offentlig telefonitjeneste for å kommunisere seg imellom. En slik utvikling vil bli ytterligere forsterket ved en innføring av Sluttbruker ENUM siden dette også kan oppfattes som et supplement til tradisjonelle telefonitjenester. Fordelen med Sluttbruker ENUM i forhold til løsninger som Skype er at E.164 nummeret kan brukes som kontaktinformasjon.

Mobiltelefoni og tradisjonell fasttelefoni er derfor under angrep fra to ulike hold. På den ene siden fra IP-telefoni som tilbyr billigere telefonisamtaler og på den andre siden fra Internetttelefoni som tilbyr enda billigere samtaler i tillegg til multimediatjenster. I tillegg er det et som sagt et voksende marked for trådløse nettverk slik at IP-telefoni i større grad kan realiseres som en mobiltjeneste.

De dominerende telekommunikasjonstilbyderne er derfor nødt til å fornye seg for å møte denne konkurransen. IMS er starten på denne fornyelsen som ofte refereres til som neste

generasjons nettverk(NGN). En ønsker å tilby tradisjonelle telefonikunder et utvidet multimedia tjenestetilbud. En av grunnene til dette er at en regner med at det å kun tilby telefoni etter hvert ikke vil være et godt nok argument for å få kunder å abonnere på en tjeneste. Dette innebærer en gradvis overgang fra linjesvitsjede nettverk til IP-baserte nettverk innenfor alle former for telefoni.

Selve IMS-arkitekturen er kompleks og det er utarbeidet flerfoldige dokumenter med spesifikasjoner. I kapitel fem i denne oppgaven ble det gjennomgått noen elementer i forbindelse med IMS, men dette var langt fra noen komplett beskrivelse av arkitekturen. Når telefonoperatørene skal ta i bruk IMS er en derfor nødt til å investere mye i en utvidelse av dagens nettverk.

Signaleringen i IMS benytter SIP som signaleringsprotokoll mellom en del sentrale nettverkselementer. Til tross for dette kan ikke IMS direkte sammenlignes med den forholdsvis enkle arkitekturen vi kjenner for IP-telefoni basert på SIP. Utviklerne av IMS bygger opp et stort og tungt rammverk med utgangspunkt i SIP. Det er allikevel ikke gitt at de ulike tilbyderne som ønsker å ta i bruk IMS er nødt til å benytte seg av alle elementene i arkitekturen.

Et avgjørende punkt for om IMS vil bli en suksess er hvorvidt tilbyderne klarer å tilby kundene nye tjenester som oppfattes som verdifulle.

8.6 Konklusjon

ENUM er utvilsomt et nyttig verktøy i forbindelse med IP-telefoni. Ved innføring av Sluttbruker ENUM vil bedrifter og organisasjoner i større grad kunne sett opp sine egne telefonitjenester uavhengig av de ulike tilbyderne. Løsninger basert på kun Sluttbruker ENUM vil ikke bli godkjent som en offentlig telefonitjeneste siden dette ikke er en alle-til-alle tjeneste. Det er også vanskelig å se for seg at en slik løsning vil kunne tilby gode løsninger for nødanrop og kommunikasjonskontroll. Ved å offentliggjøre informasjon i DNS på denne måten vil en kunne oppleve økte problemer i forbindelse med SPIT og ulike angrep i tilegg til generelle sikkerhetsutfordringer forbundet med IP-telefoni..

Ved inføring av Infrastruktur ENUM elller Privat ENUM vil bildet være noe annerledes. Her vil administrasjon av numrene ligge hos tilbyderne og konsekvensene vil først og fremst være sparte kostnader i forbindelse med ruting, og tilsvarende tapte inntekter for de tilbyderne som kontrollerer PSTN-nettet. Uavhengig av hvilke form for ENUM som benyttes vil uansett dagens PSTN-nett bli gradvis mer overflødig.

IP-telefoni over trådløse nettverk er også et område i vekst. Med innføringen av 802.11i rammeverket vil en oppnå en større grad av sikkerhet også i forbindelse med slike løsninger. Denne formen for IP-telefoni vil allikevel ikke bli noen reel konkurrent til mobiltelefoni før en får til mekanismer som sørger for en sømløs handover mellom de trådløse nettverkene uten å gå omveien via GSM/UMTS.

For å møte den økte konkurransen vil de dominerende telefoniaktørene utvikle sine egne løsninger i form av neste generasjons nettverk. Fordelene ved slike typer løsninger er at kundene skal kunne tilbys garantert tjenestekvalitet og velprøvde mekanismer for sikkerhet videreført fra dagens mobiltelefoni. IMS vil derfor fra en kundes synspunkt oppleves som en trygg og stabil løsning. Det gjenstår fortsatt å se i hvor stor grad IMS kommer til å tillate at uanhengige tredjeparts løsninger for applikasjonstjenester, og hvordan kunden skal belastes for bruken av disse.

9 Referanser

[1] Alan B. Johnston, SIP: Understanding the Session Initiation Protocol,
Second Edition, Artech House Books, 2003

[2] James F. Kurose, Keith W. Ross, Computer Networking,
A top-down approach featuring the Internet, Addison Wesley Longman Inc., 2001

[3] William Stallings, Network security essentials, Applications and Standards,
Second Edition, Pearson Education, 2003

[4] James F. Ransome, John W. Rittinghouse, VoIP Security,
Elsevier Science & Technology Books, 2004

[5] Daniel Wong, Wireless Internet Telecommunications, Artech House Publishers, 2004

[6] Miikka Poikselka, Georg Mayer, Hisham Khartabil, Aki Niemi, The IMS:
IP Multimedia Concepts and Services in the Mobile Domain, John Wiley & Sons, 2004

[7] Mathew Stafford, Signaling and Switching for Packet Telephony,
Artech House Publishers, 2004

[8] Frank Ohrtman, Voice Over 802.11, Artech House Publishers, 2004

[9] Thomas Porter, Jan Kanclirz Jr., Practical VoIP Security, Syngress, 2006

[10] David Kristensveen, Sikkerhetsutfordringer ved IP-telefoni,
Norges Teknisk-Naturvitenskaplige Universitet, 2005

[11] Internet Engineering Taskforce(IETF), SIP: Session Initation Protcol,
http://www.rfc-archive.org/getrfc.php?rfc=3261

[12] Security Group Zürcher Hochschule Winterthur: SIP Security
http://security.zhwin.ch/DFN_SIP.pdf

[13] Session Peering for Multimedia Interconnect (speermint)
http://www.ietf.org/html.charters/speermint-charter.html

[14] Internet Engineering Taskforce(IETF),Threat Analysis of the Domain Name System,
http://www.rfc-archive.org/getrfc.php?rfc=3833

[15] DNS Security Extensions, Securing the Domain Name System
http://www.dnssec.net/

[16] UMA Technology, Extending Mobile Services to Unlicensed Spectrum
http://www.umatechnology.org/

[17] UMA Technology, UMA spesifications
http://www.umatechnology.org/specifications/index.htm
[18] 3GPP, TS 23.228 IMS Stage 2 (Release 5)
http://www.arib.or.jp/IMT-2000/V440Mar05/2_T63/ARIB-STD-T63/Rel5/23/A23228-5d0.pdf
[19] A. Johnston, H. Sinnreich, SIP, P2P, and Internet Communications,
http://www.ietf.org/internet-drafts/draft-johnston-sipping-p2p-ipcom-02.txt
[20] Internet Engineering Taskforce(IETF), Secure DNS Dynamic Update
http://www.rfc-archive.org/getrfc.php?rfc=3007
[21] Avaya, Enterprise SIP Trunking
http://www.avaya.com/master-usa/en-us/resource/assets/whitepapers/lb2749.pdf
[22] Minimum requirements for interoperability of ENUM implementations
http://www.dbai.tuwien.ac.at/proj/semnum/documents/ts_102172v010201p.pdf
[23] ENUM scenarios for user and infrastructure ENUM
http://www.dbai.tuwien.ac.at/proj/semnum/documents/tr_102055v010101p.pdf
[24] ETSI, ENUM Administration in Europe
http://www.pts.se/Archive/Documents/SE/ENUM%20Administration%20in%20Europe-ETSI.pdf
[25] G. Kambourakis, D. Geneiatakis, S. Gritzalis, T. Dagiuklas, C. Lambrinoudakis,
Security and Privacy issues towards ENUM protocol
http://www.snocer.org/Paper/IT_472_CAMERA_READY.pdf
[26] Internet Engineering Taskforce(IETF), Extensible Provisioning Protocol (EPP)
http://www.rfc-archive.org/getrfc.php?rfc=3730
[27] Internet Engineering Taskforce(IETF), (EPP) Transport Over TCP
http://www.rfc-archive.org/getrfc.php?rfc=3734
[28] Internet Engineering Taskforce(IETF), (3GPP) Release 5 Requirements on SIP
http://www.rfc-archive.org/getrfc.php?rfc=4083
[29] PROTOS Test-Suite: Session Initiation Protocol(SIP)
http://www.ee.oulu.fi/research/ouspg/protos/testing/c07/sip/index.html#usage-external
[30] Avaya, Enterprising with SIP — A Technology Overview
http://www.avaya.com/master-usa/en-us/resource/assets/whitepapers/lb2343.pdf

[31] Norid, ENUM Registrarkontrakt,
http://www.norid.no/enum/registrar-kontrakt.html
[32]Internet Engineering Taskforce(IETF): Session Initiation Protocol for Telephones(SIP-T): Context and Architectures
http://www.rfc-archive.org/getrfc.php?rfc=3372
[33] Kayote Networks, SPIT prevention Security Model,
http://www.kayote.com/web/docs/WhitePapers/KayoteNetworksWhitePaper-SPIT_Prevention_Security_Model.pdf
[34] Dongwook Shin and Choon Shim, Voice Spam Control with Gray Leveling
http://www.vopsecurity.org/Qovia_Voice_Spam_control_algorithm-VoIPSecurityWorkshop_5B1_5D.pdf
[35]SPEERMINT Requirements and Terminology
http://www.ietf.org/internet-drafts/draft-ietf-speermint-reqs-and-terminology-01.txt
[36] Real Time Protocol Header
http://java.sun.com/products/java-media/jmf/2.1.1/guide/images/RTPRealTime2.gif
[37] Internet Engineering Taskforce(IETF), Number Portability in the Global Switched Telephone Network (GSTN),
http://www.rfc-archive.org/getrfc.php?rfc=3482
[38] The Internet2 SIP.edu Initiative
http://www.internet2.edu/sip.edu/docs/sip.edu-whitepaper1.pdf
[39] “Packet-Based Multimedia Communications Systems,”
ITU Recommendation H.323, 2000.
[40] Performance Technologies, SS7 Tutorial,
http://www.pt.com/tutorials/ss7/
[41] National Institute of Standards and Technology (NIST): Security Considerations for Voice Over IP Systems,
http://csrc.nist.gov/publications/nistpubs/800-58/SP800-58-final.pdf
[42] Internet Enigneering Taskforce(IETF), Media Gateway Control Protocol (MGCP)
http://www.rfc-archive.org/getrfc.php?rfc=3435

[43] Internet Engineering Taskforce(IETF), STUN - Simple Traversal of User Datagram Protocol (UDP) Through Network Address Translators (NATs)
http://www.rfc-archive.org/getrfc.php?rfc=3489
[44] Internet Engineering Taskforce(IETF), Interactive Connectivity Establishment(ICE),
http://www.croczilla.com/zap/rfcs/draft-ietf-mmusic-ice-06.txt
[45] Internet Engineering Taskforce(IETF), A Framework for TRIP,
http://www.rfc-archive.org/getrfc.php?rfc=3489
[46] Jon Hardwick, Session Border Controllers, Enabling the VoIP revolution,
Data Connection Ltd
http://www.terena.nl/mail-archives/tf-vvc/pdfYfbV3OWXf3.pdf
[47] Communication Networks Laboratory, Department of Informatics and Telecommunications, University of Athens, Security in 3G Networks
http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6TYP-4B8X2JP-4&_coverDate=05%2F01%2F2004&_alid=412250721&_rdoc=1&_fmt=&_orig=search&_qd=1&_cdi=5624&_sort=d&view=c&_acct=C000030078&_version=1&_urlVersion=0&_userid=586462&md5=aef34ee3b0770202e13031cb9ace7683
[48] Internet Engineering Taskforce(IETF), IP Encapsulating Security Payload (ESP),
http://www.rfc-archive.org/getrfc.php?rfc=2406
[49] Internet Engineering Taskforce(IETF), Internet Key Exchange (IKEv2) Protocol ,
http://www.rfc-archive.org/getrfc.php?rfc=4306
[50]Juniper Networks, Evolution of Session Border Controllers
http://www.juniper.net/solutions/literature/white_papers/200119.pdf
[51]] Internet Engineering Taskforce(IETF), The Use of RSVP with IETF Integrated Services
http://www.rfc-archive.org/getrfc.php?rfc=2210
[52] Internet Engineering Taskforce(IETF), RSVP Security Properties
http://www.rfc-archive.org/getrfc.php?rfc=4230
[53] S. Lind og P. Pfautz, Infrastrucure ENUM Requirements
http://www.ietf.cnri.reston.va.us/internet-drafts/draft-ietf-enum-infrastructure-enum-reqs-02.txt

[54]Internet-Draft, Requirements from SIP Session Border Control Deployments, http://ietfreport.isoc.org/idref/draft-camarillo-sipping-sbc-funcs/