Handover i Mobil IP

i Fremtidens Mobile Internett

Hovedoppgave
ved
sivilingeniørutdanning i
informasjons- og kommunikasjonsteknologi

av
Bjørn Badowski

Grimstad, Juni 1999

Innholdsfortegnelse

1 Forord

Denne rapporten inneholder resultater fra hovedoppgaven utført vårsemesteret 1999 av Bjørn Badowski som er student ved sivilingeniørstudiet i informasjon og kommunikasjonsteknologi ved Høgskolen i Agder.

De siste årene har vi hatt en eksplosiv vekst når det gjelder internetteknologi og mobilkommunikasjon. Salget av bærbare pc'er og små håndholdte maskiner øker, og i framtiden vil sannsynligvis bærbart utstyr utgjøre majoriteten av datamaskiner tilknyttet Internett. For å kunne fungere i mobil sammenheng må disse mobile enhetene kunne håndtere både faste og trådløse tilknytninger, uten at brukeren må behøve å gjøre noe.

Denne hovedoppgaven tar utgangspunkt i en protokoll som er designet innenfor IETF for håndtering av mobile brukere kalt Mobil IP, og oppgaven er gitt av Geir Egeland ved Telenor FoU - Kjeller.

Jeg vil først og fremst takke veileder og faglærer Geir Egeland (Telenor FoU - Kjeller) for oppgaven, god veiledning og oppfølging, samt for mange gode ideer og råd underveis.

Oslo, 28 Mai 1999

Bjørn Badowski

2 Sammendrag

Denne rapporten omhandler en utredning av handover i Mobil IP og hvordan den kan forbedres slik at den er mer egnet til framtidens mobile Internett. Begrepet handover omfatter situasjonen som oppstår når en mobilnode flytter seg fra et sted til et annet, og må bytte tilknytning for å fortsatt kunne ha kontakt med nettverket.

Rapporten begynner med en innføring av de aktuelle teknologiene som er viktige for gjennomføring av oppgaven. De viktigste teknologiene er protokollen Mobil IP, og neste generasjons Internett (IPv6).
Mobil IP er en protokoll som bygger på Internett Protokollen (IP) og gjør mobilitet transparent for applikasjoner og høyere lags protokoller. Funksjonaliteten i mobil IP går ut på at pakker som blir sendt til mobilnodens adresse i hjemmenettet vil bli tatt hånd om av en stedfortreder (en mobilitetsagent) når mobilnoden befinner seg ute, og som videresender dataene til mobilnodens nåværende lokasjon.
Videre i rapporten finnes det også beskrivelse av enkelte teknologier som er aktuelle for anvendelse i sammenheng med Mobil IP. Disse er: GPRS (i GSM), UMTS, og trådløse LAN som IEEE 802.11 og HIPERLAN.

Problemet med Mobil IP slik den er i dag er at handover fører til et kortvarig avbrudd i konnektiviteten, og pakketap. Dette vil ha stor betydning for ytelsen hvis handover forekommer ofte, noe som ikke er usannsynlig i fremtiden hvor det i større grad vil benyttes høyhastighets trådløse nett. Disse nettene som gir ytelser i Mbit/s området har liten rekkevidde, og ligger i området fra noen få meter, og opp til noen hundre meter. I en slik situasjon vil handover mellom forskjellige nett og teknologier kunne forekomme ofte, og disse avbruddene vil føre til dårligere ytelse for enkelte høyerelags protokoller, og kan føre til at såkalte tidsbundne tjenester blir tilnærmet ubrukelige.

For å forbedre handover kan det implementeres i Mobil IP forbedringer som gjør at handover kan foretas raskere, som går ut på å benytte linklagsinformasjon, bruk av *agent discovery* meldinger, og organisering av teknologier i et hierarki ut i fra deres rekkevidde og båndbredde.
I tillegg kan det også benyttes en løsning som går ut på å bygge opp nettene med mobilitetsagenter i en hierarkisk struktur, slik at det blir mulig å skille mellom forskjellige typer handover. Med en slik løsning vil mobilnoders bevegelser innenfor et administrativt domene ikke kreve noen global oppdatering, men bare internt i domenet. Fordelen med en slik løsning er at handoverytelsen vil forbedres når mobilitetsagentene befinner seg nærmere mobilnoden fordi forsinkelsen blir kortere når avstanden minskes. En annen fordel er at Mobil IP vil skalere bedre i større sammenheng, og trafikkbelastningen forårsaket av Mobil IP meldinger over Internett blir mindre.

Med disse forbedringene vil Mobil IP kunne ha et bra potensiale til å kunne bli en framtidig plattform for mobilitet i Internett forutsatt at også andre viktige områder blir forbedret (blant annet AAA tjenester, skalering, og QoS). Et eksempel på en mulig arkitektur/scenario for anvendelsen av Mobil IP er også vist i rapporten.

3 Innledning

Det siste semesteret ved sivilingeniørstudiet ved HiA består av en avsluttende oppgave kalt hovedoppgave. Oppgaven er gitt av skolen eller næringslivet, og er et prosjekt som studentene skal gjennomføre selvstendig i denne perioden. Omfanget på oppgaven er 10 vekttall, tilsvarende studiemengden normert til et semester, og har fagkoden IT6401.

3.1 Beskrivelse av oppgaven

Hovedtemaet i oppgaven er handover i Mobil IP [7], og hvordan Mobil IP kan brukes i fremtidens mobile Internett. Mobil IP er en protokoll som bygger på Internett Protokollen (IP), og gjør mobilitet transparent for applikasjoner og høyere lags protokoller som TCP.

Oppgaven går ut på å primært se på hvordan handover blir gjort i Mobil IP og kunne forslå forbedringer på dette området. Handover er situasjonen som oppstår når en mobil node forandrer sin fysiske lokasjon fra et sted til et annet, og må bytte nettverkstilknytning. Jeg vil også se noe på hvordan Mobil IP kan brukes i stor skala (skaleringsmuligheter).
I oppgaven skal jeg også undersøke hvordan funksjonalitet i IPv6 som anycast/multicast kan benyttes i forbindelse med handover.
I Europa og i Telenor vurderer man hvorvidt det er mulig å ta i bruk Mobil IP standarden som mobilhåndteringssystem til GPRS (*General Packet Radio Service*), UMTS (*Universal Mobile Telecommunications System*) og trådløse LAN. Oppgaven blir å undersøke hvorvidt dette er mulig og foreslå en egnet arkitektur for Mobil IP i forbindelse med mobile systemer.

Oppgaven er utført ved Prosjekt I hos Telenor FoU på Kjeller. Ekstern veileder for oppgaven er Geir Egeland.

3.2 Motiv

Motivet for denne hovedoppgaven er at handover vil ha stor betydning for ytelsen og tjenestekvaliteten for mobile brukere. Dette gjelder særlig i situasjoner hvor en ofte bytter mellom forskjellige nett. Derfor er det viktig at skjer så fort som mulig, slik at ikke data går tapt og dataoverføringen forringes.
Når data går tapt i en handover vil ytelsen til den mest brukte internettprokollen for ende-til-ende kommunikasjon – TCP bli dårligere fordi den regulerer datahastigheten ut i fra pakketapet og RTT (Round Trip Time). Dette vil i mobile systemer kunne gi dårlig ytelse i situasjoner hvor en ofte har handover. Handoverprosessen i Mobil IP er tidkrevende, og gjør at det blir et opphold i dataoverføringen hvor data går tapt. TCP vil dermed tro at det er metning i nettet når handover skjer, og sette ned hastigheten unødvendig.
Forsinkelsen forårsaket av handover kan også føre til problemer på et annet område som er såkalte realtime applikasjoner, hvor en er avhengig av liten eller relativ konstant forsinkelse i nettet for at applikasjonen skal kunne fungere tilfredsstillende. Et eksempel på dette er toveis tale kommunikasjonsprogrammer (kalles også VoIP – Voice over IP).
Et annet aspekt som kan få betydning for ytelsen i fremtiden er mengden av trafikk på nettet forårsaket av handover. Ettersom en regner med at en stor del av brukerne i fremtiden vil være mobile brukere, vil en betydelig del av trafikkmengden i Internett være handovermeldinger, noe som for brukeren fører til økte kostnader og unødvendig nettbelastning. Det er derfor

ønskelig å minimalisere dette mest mulig, og muligens omorganisere Mobil IP på en måte som gjør at den skalerer bedre i større sammenhenger.

3.3 Avgrensning av oppgaven

Et viktig område som jeg ikke sett på i denne oppgaven er sikkerhetsaspektet i Mobil IP. Dette området som kalles AAA tjenester (*Authentication, Accounting & Authorization*) er noe som en er helt avhengig av for å kunne benytte Mobil IP i det hele tatt.
For at Mobil IP skal kunne brukes i stor skala må det finnes mekanismer for autentisering av brukere slik at ingen er i stand til å gi seg ut for å være en annen, og ta over andre noders trafikk. I tillegg til dette må det også finnes mekanismer som håndterer dette med autorisering og betaling for bruken av tjenester. Dette vil bli mer komplekst enn det er i dag hvor systemene er homogene, og brukeren har ett abonnement for hver teknologi (ett for GSM, ISDN osv.). Med Mobil IP og mobile brukere vil det være behov for nye løsninger for håndtering av dette da brukerne vil besøke forskjellige nett, som er eid av forskjellige leverandører, som benytter seg av forskjellig type teknologier, og som tilbyr forskjellige tjenester.
Grunnen til at jeg ikke har sett nærmere på dette er at dette området alene er så stort og omfattende at det i seg selv dekker en hovedoppgave. Derfor har jeg videre i oppgaven bare gått ut i fra at dette er i orden og fungerer, eller at en for eksempel tenker seg at den mobile brukeren bare beveger seg innenfor en leverandør eller organisasjons forskjellige nett, og at brukeren er autentisert hos denne leverandøren/organisasjonen for å kunne benytte seg av disse nettene.
Derfor ser jeg bort i fra dette og fokuserer på selve handoversituasjonen som er kjernen for denne hovedoppgaven.

3.4 Mål

Utformingen av konkrete mål for hovedoppgaven har vært en dynamisk prosess. Mål og målformuleringer har forandret seg underveis, etterhvert som jeg har arbeidet med oppgaven.

Mål for oppgaven er:

- Studere og forstå hvordan teknologien fungerer, noe som innebærer at en må sette seg inn i protokollene: IPv4, IPv6, Mobil IPv4, Mobil IPv6,Trådløse LAN samt GPRS.
- Se spesielt på hvordan handover foregår i Mobil IP.
- Se på om anycast i IPv6 kan forbedre handover.
- Kunne foreslå forbedringer slik at handover kan foregå raskere, og med mindre pakketap.
- Kunne foreslå forbedringer som gjør at Mobil IP skalerer bedre i større sammenhenger.
- Se om Mobil IP er egnet i sammenheng med mobile systemer som GPRS, og framtidens mobile systemer som UMTS.

3.5 Metode

3.5.1 Fundament

Fra tidlig i januar, og fram til ca. midten av mars hadde jeg et litteraturstudie hvor mye av tiden gikk med på å sette seg inn i de aktuelle teknologiene, og forstå disse. En stor del av stoffet er modningsstoff, hvor det tar en del tid før en forstår alle mekanismer og sammenhenger.
På forhånd hadde jeg ingen spesielle kunnskaper på området, unntatt en del om IPv4, samt noe om IPv6 (fra faget datakommunikasjon IT2200), litt generelt om mobil telekommunikasjon (fra faget telekommunikasjonssystemer IT2300).
Jeg begynte først med å se generelt på Mobil IPv4 [25], og IPv6 [2,3,4]. Videre så jeg på Mobil IPv6 [9], samt dypere på Mobil IP når det gjelder selve handover situasjonen (for begge versjoner).
Jeg leste også noe om GPRS i GSM [21,22], UMTS [23], og en del om trådløse LAN teknologier [16,17,18].

3.5.2 Videre prosess

Etter hvert som jeg forstod mer og fikk en viss oversikt over emnet, prøvde jeg å samle trådene, og strukturere disse slik at en er bedre i stand til å løse oppgaven.
Det er også på denne måten rapporten er bygd opp, med et fundament først med hvor jeg fikk et generelt forståelse og overblikk over de aktuelle teknologiene. Deretter studerte jeg de delområdene som er spesielt viktige for oppgaven, og prøvde å sette dette sammen på en hensiktsmessig måte, for så å kunne benytte denne informasjonen og forståelsen for fagfeltet som jeg har oppnådd gjennom denne perioden med arbeid for å kunne foreslå forbedringer.

3.5.3 Møter

Gjennom hele prosjektperioden hadde jeg møter med veileder Geir Egeland (faste møter annenhver uke, samt en del uformelle møter til forskjellige tidspunkter etter behov), hvor vi diskuterte temaer som jeg hadde undersøkt, videre arbeide, samt diskusjon av forslag til forbedringer.

3.6 Rapportens struktur

Rapporten begynner med en innledning med oppgavedefinisjon, motiver, avgrensninger, og mål for oppgaven. Deretter følger en generell innføring til de aktuelle teknologiene som ligger til grunn for å kunne gjennomføre og forstå oppgaven.

Disse teknologiene er:
- Internett/IPv4.
- IPv6.
- Mobil IPv4.
- Mobil IPv6.
- GPRS i GSM og UMTS.
- Trådløse LAN (IEEE 802.11 & HIPERLAN).

Videre følger en mer spesifikk beskrivelse av de aktuelle emnene innenfor hver teknologi som angår hovedoppgaven spesielt:

- Mobil IP – Handover prosedyren og problemstilling
- Vertikale handover (handover mellom forskjellige teknologier).
- Bruk av anycast i IPv6 for handover.

Ut i fra kunnskaper oppnådd gjennom studie av disse emnene samt diskusjoner med veileder Har jeg kommet med forslag til forbedringer på forskjellige områder:

- Vertikale handover.
- Mobil IP handover.
- Skalering.

I tillegg finnes det i rapporten forslag på:

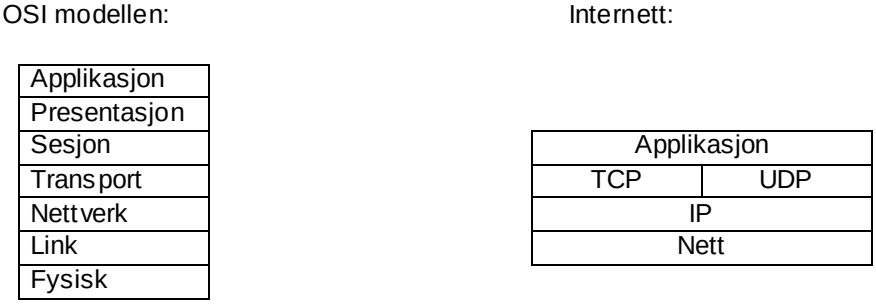
- Bruk av Mobil IP i GPRS.
- Anvendelse/arkitektur med Mobil IP.

Oppgaven avsluttes og oppsummeres med en diskusjon og konklusjon, samt helt til slutt et stikkordregister og litteraturreferanser med oversikt over litteratur som er brukt i gjennomføring av oppgaven, og som er referert til gjennom hele rapporten.

4 Innledning til teknologi – Internett - IP

Det siste tiåret har det vært en eksplosiv utvikling når det gjelder utbredelsen av Internett og internettjenester. Internett har gitt oss muligheter for innhenting og utveksling av informasjon på måter som for noen få år tilbake var umulig. Utbredelsen av WWW (*World Wide Web*), og andre tjenester som elektronisk post, filoverføring, og audio/video overføringer har gjort sitt inntog stort sett overalt: i utdanningsinstitusjoner, næringsliv, statlige institusjoner, og private husstander.
Internett har en oppbygning med høy grad av heterogenitet, hvor endesystemene i nettet varierer fra kraftige servere og arbeidsstasjoner til små håndholdte enheter.
Internettforbindelsene varierer fra høyhastighets fiberoptiske linjer med gigabit båndbredde til lavhastighetslinjer i oppringte samband med hastigheter i kilobit/s området.
På tross av denne mangfoldigheten kan noder tilknyttet Internett kommunisere med hverandre uavhengig av hva slags underliggende teknologi som benyttes i endesystemet eller dataforbindelsen. For at alt dette skal kunne spille sammen og kommunisere med hverandre på et felles språk er det Internett Protokollen (IP) som er selve kjernen.
Forenklet sett kan en se på IP som to deler: Et globalt adresseringssystem for å kunne identifisere alle nodene tilkoblet Internett, og hvor levering av data er pakkebasert. Denne leveringsmåten fungerer etter "*best effort*" metoden, som betyr at nettverket gjør så godt det kan når det gjelder levering av pakker, men gir ingen garantier om at pakken kommer frem[1].

For å kunne sikre seg at data kommer riktig frem, kunne bruke forskjellige tjenester og teknologier, er protokollene bygd opp lagvis, slik at en protokoll blir lagt over en annen. En av de første til å definere denne typen arkitektur som en felles måte til å koble noder til hverandre ble definert av ISO (*International Organization for Standardization*), og ble kalt OSI modellen (*Open Systems Interconnection reference model*)[1]. Dette er en referanse modell, og definerer ingen bestemte protokoller i hvert lag. Internett arkitekturen er også bygd opp på samme måte som OSI modellen, om enn med noe færre lag. Denne oppbygningen blir vist i figur 4-1:



Figur 4-1 - OSI modellen + Internett arkitekturen

Hvis en ser på internettarkitekturen (figur 4-1) så har en i bunnen det fysiske nettverket, som kan være Ethernet, Wireless LAN, ATM, ISDN osv. og består av den fysiske nettverksadapteren, og programvaredriveren for adapteren.
På neste nivå (nettverkslaget) finnes IP protokollen som tilbyr sin *best effort* tjeneste når det gjelder levering, og global adressering av noder.
Laget over er transportlaget hvor protokollene TCP (*Transmission Control Protocol*) og UDP (*User Datagram Protocol*) finnes. TCP gir brukeren en pålitelig tjeneste, ved å kontrollere at alle pakker har kommet frem, sørge for riktig rekkefølge, kontrollere etter feil, be om retransmisjon av pakker og benytte andre mekanismer for å sørge for at data kommer riktig frem med best mulig ytelse. UDP er en enklere protokoll som tilbyr en upålitelig tjeneste, og

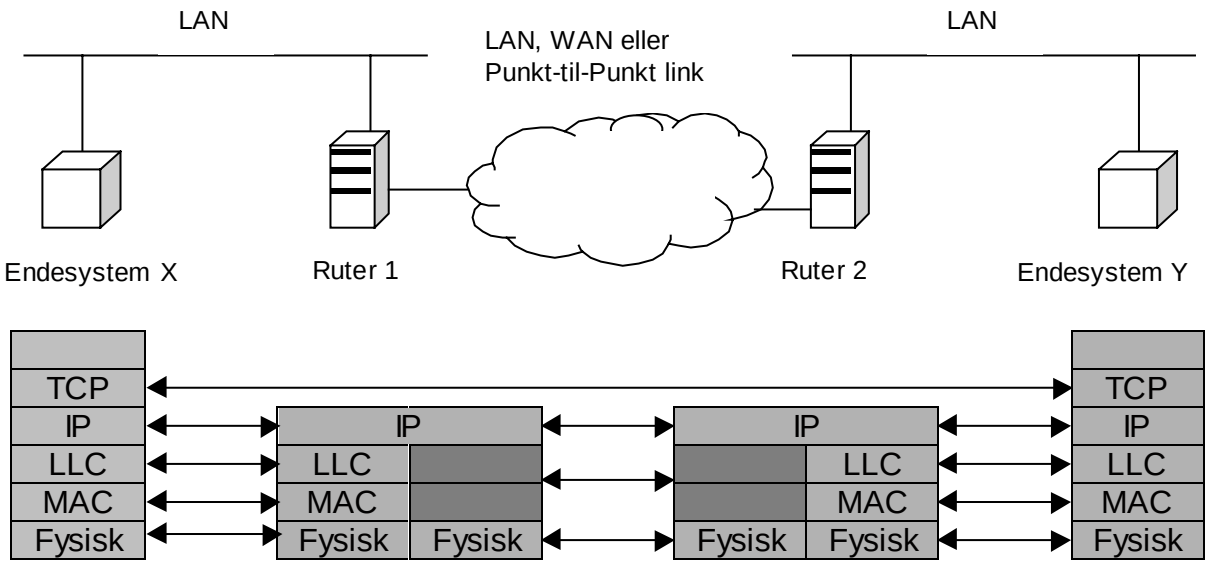
har mindre overhead enn TCP ettersom en mindre andel av hver pakke brukes til headeren da den ikke har noen mekanismer som beskytter mot feil. Dette gjør at denne er bedre egnet i enkelte situasjoner.
I det øverste laget finnes applikasjonene som benytter seg av de underliggende protokollene. Disse kan være FTP (Filoverføring), Telnet (Fjern innlogging), SMTP (e-mail), og HTTP (World Wide Web), som benytter seg av TCP for sikker overføring av data. Over UDP har en for eksempel applikasjoner for sending og mottak av audio/video og DHCP (En protokoll som tilbyr automatisk konfigurering av IP adresser for noder).

4.1 En internettprotokolls rolle

IP tilbyr funksjonalitet for sammenkobling av endesystemer over mange og flere forskjellige typer nett. For dette formålet er IP implementert i endesystemer, og rutere (som er enheter som kobler sammen og håndterer trafikk mellom flere nett). Data fra applikasjoner og høyere lags protokoller i endesystemene blir pakket inn i IP pakker for sending. Disse pakkene blir da sendt over ett eller flere nettverk, gjennom en eller flere rutere for å nå destinasjonen. Ruterne og endesystemene må derfor ha en rekke egenskaper for å kunne håndtere dette:

- **Addressesystemer** – De underliggende nettene kan være basert på forskjellige arkitekturer. For eksempel skal en ruter håndtere trafikk fra et Ethernet LAN med 48-bits hardware adresser, som er tilkoblet et ATM nettverk eller en X.25 tilkobling. For at dette skal fungere må en ha en form for global adressestruktur.
- **Forskjellige pakkestørrelser** – Pakker sendt fra en type nett må kanskje deles opp for å kunne sendes på andre typer nett, noe som kalles fragmentering. For eksempel så har ethernetpakker en maksimum nyttelast (*payload*) på ca. 1500 bytes, mens på et X.25 nettverk er 1000 bytes vanlig. En pakke sendt fra et ethernet som skal bli videresendt av en ruter over et X.25 nett må derfor dele opp pakken i to mindre deler for å kunne sende den videre.
- **Grensesnitt** – Hardware og software grensesnittene mellom forskjellige nett varierer, og ruteren må kunne håndtere dette.
- **Pålitelighet** – Forskjellige nett kan tilby både pålitelige ende til ende oppkoblinger, og upålitelige tjenester. Dette må en endesystemene kunne håndtere, og må derfor være uavhengig av andre netts pålitelighet.

Som eksempel på hvordan dette kan foregå kan en se på figur 4-2.



Figur 4-2 – Nettverkslag og rutere - illustrasjon

Figur 4-2 viser en situasjon hvor to endesystemer kommuniserer med hverandre med mellomliggende rutere vedå bruke IP. Endesystemene må ha protokoller i alle lag helt opp til applikasjonslaget, mens ruterne bare trenger protokoller opp til nettverkslaget (IP). Hvis en skal se på pakkeflyt fra endesystem X til Y, vil data fra applikasjonen (for eksempel et FTP program) pakkes inn i en TCP pakke. Denne blir så pakket inn i en IP pakke med destinasjonsadressen til Y. IP laget finner ut at Y er på et annet subnett, og gir laget under (LLC – Logical Link Control) riktig adresse informasjon (dvs. ruteren). LLC laget sender IP pakken ned til MAC laget sammen med riktig adresseinformasjon, som konstruerer en MAC (*Medium Access Control*) pakke med hardware adressen til ruteren (ruter 1), og sender pakken ut på det fysiske nettet. Pakken som blir sendt ut på lokalnettet vil se slik ut:

MAC Header	IP Header	TCP Header	Applikasjonsdata

Figur 4-3 - Pakke med headere

Når pakken blir tatt i mot av ruter 1, stripper den av headerne opp til IP nivå, og analyserer IP headeren for å finne veien til Y. Her har en to muligheter:

- Endesystem Y er direkte tilkoblet ruteren på ett av subnettene ruteren har direktekontakt med.
- Endesystem Y er på et subnett ruteren ikke har direktekontakt med, og pakken må sende gjennom en eller flere andre rutere.

I dette eksempelet må pakken rutes gjennom et mellomliggende nett til ruter 2 for å nå fram til Y. Hvis nettet i mellom ruterne er et X.25 nettverk blir IP pakken pakket inn i en X.25 pakke, og adressert med ruter 2 sin adresse. Når pakken mottas av ruter 2 blir X.25 headeren fjernet, og ruteren finner ut at destinasjonsadressen tilhører et endesystem på et subnett ruteren er direkte koblet til, og den lager da en pakke med riktig hardware (MAC) adresse for Y og sender denne til Y. Når pakken til slutt mottas av Y, vil headerne igjen strippes vekk helt opp til applikasjonslaget, og pakken har dermed nådd sin destinasjon.

Dagens IP (IPv4) kan ikke tilby noen form for tjenestekvalitet, og fungerer etter *best-effort* metoden, dvs. den gjør så godt den kan. Den garanterer ikke at pakker kommer frem, eller at de når fram i samme rekkefølge. Det er protokollaget over som har ansvar for dette (i dette tilfellet TCP), som må finne ut om pakker mangler, er blitt ødelagt, eller har fått byttet om rekkefølgen. Ved å bygge opp nettverksstrukturen på denne måten gir god fleksibilitet,

muliggjør bruk av forskjellige nettverkstyper, samt at pakker kan gå forskjellige veier gjennom nettet, noe som gjør at protokollen kan velge andre veier hvis det er mye trafikk på enkelte linker, eller at linker går ned.

5 IPv6

5.1 Innledning

Med den raske utviklingen av dagens Internett når det gjelder trafikkmengde, økning i antall brukere og dermed etterspørsel etter IP adresser, økte krav til tjenestekvalitet og sikkerhet begynner dagens IP protokoll å vise sine begrensninger.
Opprinnelig var IP beregnet for enkle distribuerte applikasjoner, filoverføring, e-mail og fjerntilgang som Telnet, men i dag har Internett blitt et multimedia og applikasjonsrikt medium basert på web teknologi. Bedrifter og skoler bruker ikke lenger nettene til e-mail og enkel filoverføring, men har gått over til bruk av komplekse klient/server miljøer, og bruk av intranettløsninger.
Alle disse utviklingene har satt høyere og høyere krav til de IP baserte nettenes ytelse og tjenester. En har fått behov for nettverk som kan håndtere sanntidstrafikk, fleksible metningskontroll systemer, og sikkerhetsfunksjoner. Ingen av disse kravene er enkle å få til med dagens IP.
Den største drivkraften bak utviklingen av en ny IP protokoll er problemet med adresserommet. Saken er den at vi går snart tom for IP-adresser. Det 32 bit lange adressefeltet er utilstrekkelig for den eksplosive veksten i Internett en har i dag, og på grunn av måten adresserommet er administrert.
På grunnlag av dette har man konstruert en ny protokoll kalt IPv6 (IP versjon 6) [2,3,4, 26], som omsider skal erstatte dagens IP (IPv4).

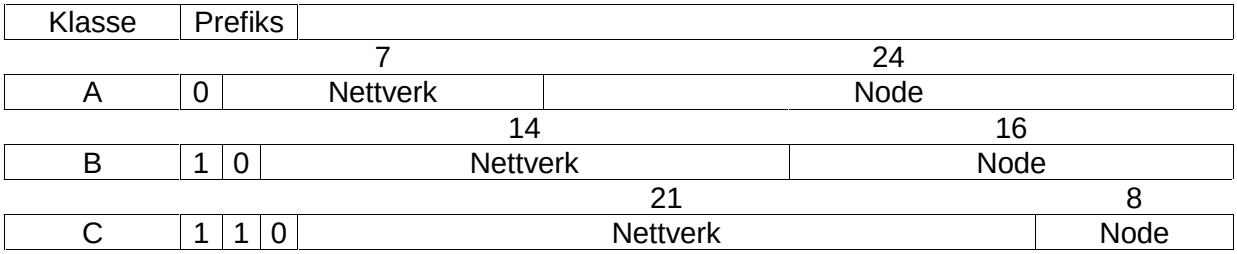
5.2 Neste generasjons IP

Etter flere år med stor vekst i trafikkmengde og antall tilknyttede noder ble det mer og mer tydelig at Internett med dagens IP (IPv4) ikke er i stand til å møte framtidens funksjonelle og ytelsesmessige krav. Følgene av dette var at IETF (*Internet Engineering Task Force*) startet utviklingen av neste generasjon Internett Protokoll – IPv6. Hoveddrivkraften bak utviklingen var det begrensede adresserommet, samt en rekke andre krav og ønsker som har oppstått. Disse kan ses på som 5 hovedområder: Adresserom, ytelse, nettverkstjenester, addresseringsfleksibilitet, og sikkerhet.

5.2.1 Adresserom

Dagens IP (v4) har et 32 bits adressefelt som teoretisk gir plass til 2^{32} forskjellige adresser dvs. over 4 milliarder mulige adresser. En skulle kunne tro at dette skulle være mer enn nok, men mot slutten av 80 åra begynte man å forstå at dette vil bli et problem i fremtiden. Noen av grunnene til dette var:

- I IPv4 benyttes en to-nivå struktur på IP adressene (nettverksnummer og nodenummer), og tre forskjellige klasser av internett adresser (A, B, og C). Klasse A bruker 8 bits for nettverket, og resten (24 bits) for nodenummeret, klasse B 16 i hver, og klasse C har 24 bit for nettverksnummeret, og 8 bits for nodenummer. De forskjellige klassene blir vist i figur 5-1.



Figur 5-1 - IPv4 adressestruktur

Denne oppdelingen er praktisk i rutingsammenheng, men fører til sløsing av IP adresser. Dette er fordi at når et nett får tildelt et nettverksnummer, for eksempel en klasse C adresse, vil den eie alle nodenumrene for dette nettet (8 bit vil gi 254 mulige noder på nettet). Hvis ikke alle adressene blir brukt går de som er til overs til spille, da de ikke kan benyttes av andre.

- Det er vanlig praksis å gi unike nettverksnummer til IP nettverk uavhengig om de er tilknyttet Internett eller ikke. Det er mulig å ha gjenbruk av IP-adresser på private nett ettersom de ikke er tilknyttet Internett, men det kan være risikabelt og tungvint hvis nettet skal kobles opp mot Internett senere.

- Internett generelt har hatt en eksplosiv vekst i mange år. Antall nettverk øker raskt, mange organisasjoner har ikke lenger bare ett LAN (*Local Area Network*) men flere. Trådløse nett har også begynt å dukke opp, og når dette virkelig tar av vil en få behov for enda flere adresser.

- Bruk av TCP/IP på nye områder vil kreve enda større behov for nye unike IP adresser. De nye bruksområdene som vil føre til behov for enda flere IP adresser er PDA'er (Personal Digital Assistant), mobiltelefoner, og muligheter for kommunikasjon med "ting som tenker" som TV, video, stereo, kjøleskap osv.

- Vanligvis har en kun en IP-adresse per node, men ved å bruke flere IP-adresser per node vil en kunne få nye muligheter, noe som krever enda flere IP-adresser.

Før å møte alle disse behovene har en i IPv6 fått et adressefelt på 128-bit istedenfor 32-bit som i IPv4. Dette er en økning i en faktor på 2^{96}. Dette burde holde lenge selv om adressene blir tildelt ueffektivt. Hvis en adresserer like ueffektivt som en gjør i dag skulle dette gi noe over 1500 adresser per kvadratmeter over hele jordas overflate! (Dette er beregnet ut fra en analyse av effektivitet i adressearkitekturer franske, og amerikanske telefonsystemer, Ethernet LAN IEEE 802.3, samt dagens Internett).

5.2.2 Ytelse

I LAN og WAN (*Wide Area Network*) nettverk har ytelsen økt betraktelig de siste årene, med hastigheter opp i hundretalls megabit per sekund, samt også muligheter for gigabit hastigheter. På Internettfronten har en fått nye tjenester som for eksempel webapplikasjoner med mye grafikk og multimedia, noe som krever større båndbredder enn de tradisjonelle tjenestene. Med denne økte kapasiteten i nettene og nettbelastningen som nye tjenester fører til er det viktig at ruterne har god ytelse. Ruterne må være i stand til å prosessere å videresende IP pakkene så fort at en får utnyttet høyhastighetsnettene, og opprettholdt trafikkflyten best mulig. Den faktoren som har størst betydning i denne sammenhengen er ruternes hardware-arkitektur, men IP pakkens design spiller også en viktig rolle. Ved å bygge

opp pakkene på en effektiv måte kan en gjøre at prosesseringen per pakke blir mindre, som igjen gjør at det blir lettere og billigere å lage rutere med høy ytelse.

IPv6 har tre elementer som er forbedret i forhold til IPv4 for å kunne gi økt ytelse:

1. I IPv6 har headeren fast lengde, mens i IPv4 har headeren variabel lengde. Dette fører til at prosesseringen av pakker blir enklere.

2. Antall felter i IPv6 headeren er redusert i forhold til IPv4. Flere av tilleggene en har i IPv6 blir lagt i tilleggsheadere som blir plassert mellom IPv6 headeren og payload (Mer info om dette i kap. 5.3). De fleste av disse tilleggsheaderne blir ikke prosessert av rutere. Dette gjør at prosesseringen av IPv6 pakker blir enklere og kan derfor gjøres raskere enn i IPv4.

3. I IPv6 har ikke ruterne lov til å fragmentere pakker, mens dette kan bli gjort i IPv4. I IPv6 kan fragmentering bare bli gjort hos avsender. Ettersom ruteren i IPv6 også får færre oppgaver, vil dette også kunne føre til økt ytelse.

5.2.3 Nettverkstjenester

I IPv6 finnes det et prioritetsfelt i headeren, som gjør at brukeren kan spesifisere ønsket tjenesteklasse for datatrafikken (QoS – *Quality of Service*). I ruterne vil ruting av pakker blir gjennomført på grunnlag av disse klassene. Dette er spesielt gunstig for real-time tjenester som krever høy eller konstant bitrate, som for eksempel videooverføring, hvor man spesifiserer at denne pakkestrømmen skal ha høy prioritet. I ruteren vil dette føre til at lavproritetspakker blir forkastet først når det høy trafikk og opphopning av pakker. IPv4 tilbyr minimal hjelp på dette området, det finnes tilleggsprotokoller for dette men er lite i bruk i dag. I IPv6 ligger dette i standarden, og blir derfor støttet av alle noder.

5.2.4 Adresseringsfleksibilitet

Dagens IP er konstruert for ”*unicast*” adresser, som betyr at en enkelt IP adresse peker kun på en bestemt node. Det er liten støtte for andre former for adressering, delvis på grunn av at adresserommet er så begrenset, og fordi ingen adresser har blitt satt for spesiell adressebruk. IPv6 har lagt til muligheten for såkalte ”*anycast*” adresser, hvor pakker som blir sendt til en slik adresse blir levert til én node i en gruppe av noder. Multicast funksjonaliteten og skalerbarheten er også forbedret ved at en har satt av et eget adresseområde for dette formålet. (For mer info om adressering i IPv6 se kap. 5.4).

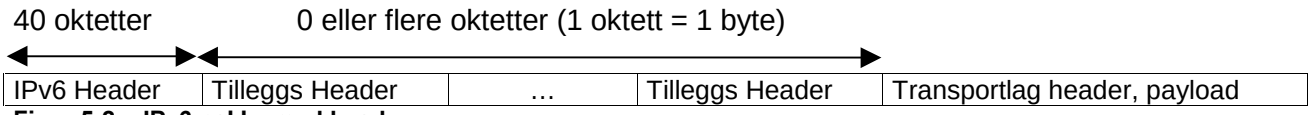
5.2.5 Sikkerhet

IPv4 har ingen innebygde sikkerhetstjenester unntatt ett valgfritt tilleggsfelt som kan legges til IP headeren. Sikkerhet kan også implementeres i høyere lags protokoller og i applikasjoner, men er mindre gunstig da en vil få løsninger som er proprietære, og lite plattformuavhengige. I IPv6 finnes det en standard sikkerhetstjeneste på IP nivå som heter IPsec [32] som enhver applikasjon kan benytte uten at en har spesielle sikkerhetstjenester lagt inn i programmet. IPsec finnes også i IPv4, men er et valgbart tillegg, og støttes derfor ikke hos alle noder. Ettersom dette er standard i IPv6, vil dette støttes av alle noder, og gi standardiserte funksjoner for viktige områder som autentisering, dataintegritet, og kryptering.

5.3 IPv6 formatet og funksjoner

5.3.1 IPv6 pakkeformat

En IPv6 pakke har den generelle formen som vist under:



Figur 5-2 - IPv6 pakke med headere

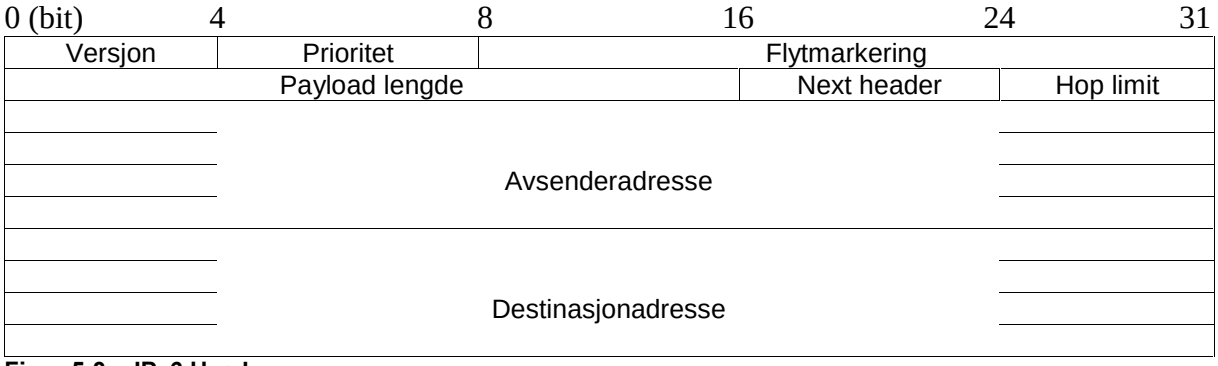
Den eneste delen av pakken som alltid må være med er IPv6 headeren som har en fast størrelse på 40 oktetter (IPv4 har 20 oktetter som minimumsstørrelse). I tillegg er det i IPv6 definert en rekke tilleggsheadere.
Disse er:

- Hop-by-hop header
- Rutingheader
- Fragmenteringsheader
- Autentiseringsheader
- Krypteringsheader (*Encapsulation security payload* header)
- Destinasjonsheader

Foran hver tilleggsheader har en også et ”neste header” felt som forteller hva slags header som er den neste tilleggsheaderen, eller hvis det er den siste tilleggsheaderen forteller den hvilken transportlag protokoll som brukes.

5.3.2 IPv6 header

IPv6 headeren er bygd opp som vist i figuren 5-3.:



Figur 5-3 - IPv6 Header

Headeren består av disse feltene:

- Versjon (4 bits): Denne forteller IP versjon nummeret, og har veriden 6.

- Prioritet (4 bits): Denne forteller om pakken bruker en av to mulige prioritetsrelaterte karakteristikker, samt 8 nivåer innenfor hver av disse. Disse to karakteristikkene er:
  - Metningskontrollert trafikk (hvor kilden for eksempel bruker TCP).

  - Trafikk uten metningskontroll (Hvor kilden ønsker konstant datarate, og/eller forsinkelse).

- Flytmarkering (24 bits): Brukes for å markere en flyt/gruppe av pakker som krever spesiell håndtering av ruterne i et nett.

- Payload lengde (16 bits): Denne forteller den resterende lengden etter headeren i ant. oktetter, dvs lengden på evt. Tilleggsheadere + transport lag PDU'en.

- Next header (8 bits): Fortelle hvilken type header som kommer rett etter IPv6 headeren.

- Hop limit (8 bits): Antall lovlige hopp for pakken. Blir satt av kilden, og minsket med en for hver ruter pakken passerer. Pakken kastes dersom verdien blir 0.

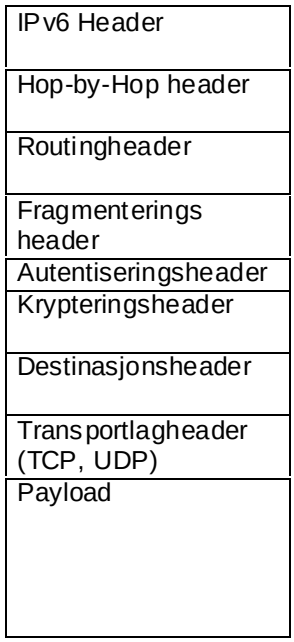
- Avsenderadresse (128 bits): Opphavsadressen til pakken.

- Destinasjonsadresse (128 bits): Adressen til den tilsiktede mottakeren av pakken. (Behøver ikke å være den endelige adressen, hvis en også har en ruting header).

Selv om denne headeren er mye større enn IPv4 headeren, inneholder den færre felter (8 mot 12), noe som gjør at prosesseringen hos rutere vil kunne gå raskere.

5.3.3 IPv6 Tilleggsheadere

En IPv6 pakke med alle tilleggsheadere vil se ut som vist i figur 5-4.



Figur 5-4 – IPv6 pakke med tilleggsheadere

Alle tillegsheadere begynner med et "neste header" felt som forteller hva slags innhold det neste feltet har. Ved å gjøre det på denne måten trenger ikke alle noder som pakken passerer å prosessere alle headere, kun den eller de som har betydning for den aktuelle noden. Hvis vi for eksempel har en router som pakken passerer gjennom vil den kun prosessere de headerne som har med ruterens oppgaver å gjøre.

Hop-by-hop headeren må undersøkes av alle rutere som pakken passerer. Hittil er bare én type tilleggsopsjon definert: nemlig mulighet for såkalt "jumbo payload". Dette muliggjør bruk av større pakker enn 65kbyte. Her finnes det et 32 bits felt i stedet, noe som muliggjør pakkestørrelser opp til 4Gbyte. Feltet legger også til informasjon hvordan en ruter som ikke skjønner meldingen skal reagere (en ruter som ikke har denne funksjonen implementert). Her kan man sette om headeren skal hoppes over, pakken forkastes, og/eller om en ICMP feilmelding skal sendes tilbake til avsenderen, for å si fra at opsjonen er ukjent.

Rutingheaderen inneholder data (dvs. adresser) om hvilke rutere pakken må passere før pakken når fram til destinasjonen. På denne måten kan det settes opp en fast rute for pakker gjennom nettet.

Fragmenteringsheaderen blir brukt til å finne ut hvor stor MTU (Message Transfer Unit) kan være på veien til en gitt destinasjon. Dette må finnes ut ettersom det bare er avsenderen som kan fragmentere pakker i IPv6.
Hvis avsenderen ikke gjør dette må den ha en maksimum størrelse på 576 oktetter på pakker som er minimumsstørrelsen som alle nett må støtte.

Autentiseringsheaderen blir brukt for at en skal kunne være sikker på at den mottatte pakken kommer fra kilden som er angitt i headeren, og at ingen data i pakken har blitt forandret underveis.

Krypteringsheaderen gir muligheter for kryptering av data for å hindre at dataene kan bli lest av tredjepart.

Destinasjonsheaderen inneholder ekstra opsjoner som destinasjonen skal prosessere. Denne er formatert på samme måte som hop-by-hop headeren. Den blir blant annet brukt i Mobil IPv6 for registreringsbeskjeder.

5.4 IPv6 Adressering

Adressene i IPv6 har et adressefelt på 128 bit, noe som muliggjør inndeling av adresserommet i et slags hierarki, basert på nettverk, nettverkstilbyder, land (geografi), firma osv. i adressen. Ved å organisere adressestrukturen på denne måten vil ruting kunne bli enklere og raskere da en trenger mindre ruting tabeller, og oppslag i disse vil bli raskere. I IPv4 er det stort sett ingen adressestruktur som gjør ruting enkelt, og ruterne må derfor ha store rutingtabeller.
En annen mulighet som finnes i IPv6 er at hvert grensesnitt på en node kan ha flere IP adresser, som gjør at en node kan bruke flere nettilbydere på et grensesnitt.
I IPv6 finnes det tre forskjellige typer i motsetning til IPv4 som kun har de to første:

1. ”*Unicast*”: En adresseidentifikator for et grensesnitt (interface). En pakke som blir sendt til en unicast adresse blir levert til grensesnittet som adressen identifiserer.

2. ”*Multicast*”: En identifikator for et sett med grensesnitt dvs. en gruppe av noder. En pakke sendt til en multicast adresse blir levert til alle nodene som tilhører multicastadressen.

3. ”*Anycast*”: En identifikator for et sett av grensesnitt (typisk flere forskjellige noder). En pakke sendt til en anycast adresse blir levert til ett av grensesnittene som adressen identifiserer basert på avstanden til nodene.

For å skille mellom de forskjellige typene adresser begynner alle med et prefiks som identifiserer de forskjellige kategoriene av adressen.

5.4.1 Unicast adresser

I IPv6 er det definert flere forskjellige unicast adresser:
- Provider based global
- Link-local
- Site-local
- IPv4 compatible IPv6
- loopback adresser

En ”provider based global” adresse gir global adressering for bruk til alle noder tilkoblet Internett. Adressen har 5 felter etter prefikset:

Prefiks

010	Registry ID	Provider ID	Subscriber ID	Subnet ID	Interface ID

Figur 5-5 – IPv6 unicast adresse oppbygning

- Registry ID: Identifiserer hvilken myndighet som tildeler provider delen av adressene
- Provider ID: En bestemt tilbyder som deler ut abonnent (subscriber) delen.
- Subscriber ID: Identifiserer forskjellige abonnenter koblet til en tilbyder.
- Subnet ID: En gruppe av noder innenfor en abonnents nett.
- Interface ID: Identifiserer en enkelt nodes grensesnitt, blant flere i et subnett.

Det er ingen fast lengde på hver av disse feltene. Et eksempel på en adresse kan være: 3FFE:2A00:1FF:F002::xxx

hvor 3 er prefikset, FFE 6bone (myndighet), 2A Norge, 001 ISP – Uninett, FFF Link, og 002 Telenor. Den siste delen av adressen bruker en til å identifisere hver node, og her kan en for eksempel bruke den 48 bits hardware adressen som ligger i vanlige nettverkskort.

”Link local” og ”site local” er adresser som en bare kan bruke lokalt i et subnett. Site-local adressene er formatert slik at de ved en senere anledning kan omgjøres til globale adresser, og er bygd opp på samme måte som provider based global adressene bortsett fra at registry, provider, og subscriber ID feltene står tomme.

Et viktig emne for implementeringen av IPv6 er overgangen fra IPv4. Det er ikke praktisk mulig å erstatte alle IPv4 rutere og oppgradere alle noder til IPv6 på en gang. En vil derfor få en periode med begge nett. Derfor er det definert et adresseområde for IPv4 kompatible IPv6 adresser. Adressen er bygd opp slik at i de 32 nederste bit er en IPv4 adresse, og hvor prefikset er satt til 96 nuller. I overgangsperioden så må alle IPv6 pakker pakkes inn i en IPv4 pakke for å kunne sende med IPv6 over nett/rutere med IPv4.

”Loopback” adressen brukes for å sende pakker til seg selv, og blir ikke sendt ut av noden. Adressen er gitt som 0:0:0:0:0:0:0:1.

5.4.2 Multicast adresser

Multicast adresser gir muligheten til å kunne adressere en forhåndsbestemt gruppe av noder med en enkelt adresse. En pakke med en multicast destinasjonsadresse blir sendt til alle medlemmene i gruppen. En multicast adresse består et fast prefiks (hex FF – som tilsvarer 1/256 av adresserommet i IPv6), samt et 4 bits flagg felt, 4 bit scope felt, og et 112 bit gruppe ID felt.

Prefiks

11111111	4 bit Flagg	Scope	Gruppe ID

Figur 5-6 – IPv6 Multicast adresse oppbygning

Flaggfeltet består av 3 nuller og et T-bit (det er ikke definert noe for de første 3 bitene ennå). T-bitet identifiserer om multicast adressen er permanent (T = 0), eller en midlertidig adresse (T = 1).

Scope feltet blir brukt til å begrense omfanget til multicast gruppen, og bestemmer om den er global for hele Internett, en site, eller et subnett (noen av verdiene er ikke i bruk ennå, og er satt av for evt. nye funksjoner/områder i fremtiden).

Gruppe ID feltet identifiserer multicast gruppen som enten er permanent eller midlertidig. Scope feltet er uavhengig av gruppe ID feltet, og brukes til å bestemme om pakker sendt til multicast gruppen kun skal mottas enkelte av medlemmene i gruppen. Et eksempel på bruk av dette er at dersom en har en gruppe servere som har en felles multicast adresse, kan scope feltetet benyttes til å kun sende pakker til servere som befinner seg i samme subnett eller samme site som avsenderen.

5.4.3 Anycast adresser

Ved å bruke anycast adresser kan en bruker spesifisere at han vil kontakte en node i en gruppe av noder som har en felles adresse. En pakke med denne typen adresser vil alltid bli rutet til den noden/grensesnittet som er nærmest i avstand fra noden (sett fra ruterens side – dvs. vanligvis antall hopp).
Oppbygning av en anycast adresse:

n bit	121-n bit	7 bit
Subnett prefiks	111111X111…11	Anycast ID
	Interface identifikator	

Figur 5-7 – IPv6 anycast adresse oppbygning

Adressen er bygd opp på samme måte som en vanlig unicast adresse bortsett fra grensesnitt ID biten. Her er det en fast struktur, hvor den siste delen som består av 7 bit identifiserer én anycast adresse i det subnettet som subnettprefikset tilsier. Det midterste feltet består av bare ettall, med unntak av bit nr. 7 (markert med X) som bestemmer om anycast adressen er lokal (0) eller ikke.
Et eksempel på bruk av anycast adresser er ved bruk av dette i en ruting header (tilleggsheader), for å spesifisere at pakken skal gå gjennom en gruppe av rutere som har en bestemt anycast adresse som en bestemt ISP (Internet Service Provider - Tilbyder), eller bestemt subnett.
Anycast adresser benytter det samme adresserommet som unicast adresser, og som kan ses ut fra oppbygningen vist i figur 5-7 er det de øverste adressene i hvert subnett som er satt av til anycast. For at en node skal bli medlem av en anycast adresse må den konfigureres for mottak på denne adressen, samt at ruterene som anycast nodene er tilknyttet til må konfigureres spesielt for at de skal koble en anycastadressen mot en gruppe av unicast adresser (dvs. hvert av medlemmene i anycast gruppen).

6 Mobil IP

6.1 Innledning

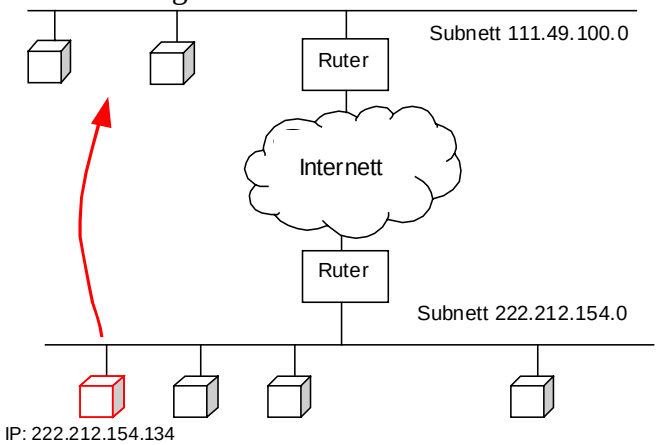
Mobil IP er en protokoll som bygger på Internett Protokollen og gjør mobilitet transparent for applikasjoner og høyerelags protokoller som for eksempel TCP. ”Mobile IP” er utviklet av en arbeidsgruppe innenfor IETF (Internet Engineering Task Force)[26,7,8]. Den vil kunne dekke behovene til mobile brukere som ønsker å ha kontinuerlig nettilgang ikke bare når en befinner seg stasjonært på forskjellige steder, men også at en opprettholder alle forbindelser mens en beveger seg fra sted til sted.

Situasjonen slik den er i dag er at vi ikke har ekte mobilitet. I beste fall gir dagens løsninger portabilitet da det er mulig å ha en eller annen form for nettilknytning hele tiden. Et eksempel på portabilitet kan være en bruker som benytter en laptop, og bruker fast nettilknytning når han befinner seg på kontoret og hjemmet, og oppringt nettilgang over GSM når han er i bevegelse. Eventuelle forbindelser som noden har med nettverket vil bli brutt når den går over fra kontornettet og over til GSM sambandet. Dette gir ikke ekte mobilitet da alle internettforbindelser blir brutt når en forflytter seg fra et tilknytningspunkt til et annet.

Det ble tidlig funnet ut at hvis en utviklet en løsning som ga ekte mobilitet på nettverkslaget (ved å modifisere IP protokollen), kunne en oppnå store fordeler som transparenthet for høyere lags protokoller og applikasjoner. At funksjonen er transparent for applikasjoner er viktig da det er uakseptabelt hvis mobile brukere må kjøpe spesielle programmer/versjoner beregnet for mobilt bruk. Avbruddsfri nettilgang er viktig for å gi høy tjenestekvalitet til brukerne, og Mobil IP er den eneste protokollen på Internett i dag som støtter dette.

6.1.1 Hvorfor er ikke mobilitet på Internett enkelt?

Internett protokollen (IP) ble opprinnelig utviklet fra en modell som gikk ut i fra at en nodes tilkobling til Internett var statisk. Som tidligere fortalt er en IP adresse bygd opp av to elementer, en identifikator for subnettet noden er tilkoblet (*Network ID*), og en unik identifikator innenfor subnettet noden befinner seg (*Host ID*). Disse to elementene er statiske, og en mobil node kan ikke ta med seg IP adressen til et annet nett uten videre.
Dette blir vist i figur 6-1:



Figur 6-1 – Internett & subnett

Hvis noden på figuren som befinner seg i subnett 222.212.154.0 og har IP adresse 222.212.154.134 flytter seg over til subnett 111.49.220.0 og beholder sin gamle IP adresse vil ingen pakker komme frem til noden lenger. Grunnen til dette er at ruting av IP pakker fungerer slik at ruteren leser destinasjonsfeltet i pakkens IP header, finner subnettnummeret, og sender pakken til det subnettet. Dermed så vil alle pakker som blir sendt til nodens IP adresse bli sendt til det gamle subnettet, og ikke der hvor noden befinner seg nå.
Ruten som pakken går gjennom nettet er heller ikke nødvendigvis den samme i begge retninger, samt at den kan forandres kontinuerlig i hver retning på grunn av nettverkets tilstand (rutere kan velge forskjellige veier avhengig av trafikkmengde og forsinkelse i nettet).

Disse to elementene er på en måte det motsatte av mobilitet. Først må noden ha en stabil (konstant) IP adresse for å enkelt kunne nås av andre enheter på Internett. Når IP adressen er fast, så er også stedet som pakkene blir rutet fram til det samme hele tiden ettersom subnettet må befinne seg på samme sted hele tiden, noe som resulterer i at det blir ingen mobilitet. For at mobilnoden skal kunne bevege seg må den derfor bytte IP adresse, noe som skaper problemer for transportprotokollen TCP. I TCP blir IP adressen, og portnummeret brukt for å identifisere forbindelser, og hvis IP adresse endres vil forbindelsen brytes.

Mobil IP spesifiserer arkitekturer og mekanismer som er designet for å løse disse problemene. Dette løses ved at mobilnoden benytter to IP adresser:

- En fast hjemmeadresse (*Home Address*) som er uavhengig av hvilket subnett den er tilkoblet, og blir brukt for identifikasjon.

- I tillegg får noden en mobil adresse (*Care Of Address*), som er midlertidig og endres for hvert tilknytningspunkt, og blir brukt til ruting til nodens nåværende lokasjon i nettet.

For at pakker som blir sendt til den faste adressen i hjemmenettet skal komme frem til mobilnoden finnes det en videresendingsmekanisme i såkalte mobilietsagenter som sørger for at pakker blir levert til mobilnodens midlertidige lokasjon i nettet. Mobilitetsagenten som er plassert i hjemmenettet vil ha rolle som mobilnodens stedfortreder når den ikke befinner seg i hjemmenettet.

Mobil IP finnes i dag i to versjoner, en for IPv4, samt en for IPv6, og videre følger en beskrivelse av disse to.

6.2 Mobil IPv4 – Oversikt

Hva er Mobil IP?

Som nevnt tidligere er Mobil IP en modifikasjon til IP protokollen. Denne modifikasjonen innebærer blant annet å legge til ekstra kontrollbeskjeder som blir benyttet i lokasjonsforandringer. Mobil IP ble spesifisert for å kunne håndtere mobile noder som beveger seg (dvs. forandrer tilknytning), og protokollen er i stand til å håndtere lokasjonsforandringer så ofte som tiden det tar for kontrollbeskjeder å bevege seg mellom hjemmenettet og mobilnoden.
Det ble satt 5 generelle krav til Mobil IP protokollen:

1. Mobilnoden må kunne kommunisere med andre noder etter linklags-tilknytningen er forandret, uten at IP adressen er forandret.

2. Mobilnoden må kunne kommunisere med andre noder som ikke har implementert Mobil IP (bakover kompatibilitet). Det skal heller ikke være behov for ekstra funksjoner i rutere, hvis de ikke skal ha roller som mobilitetsagenter.

3. Alle kontroll- og konfigurasjonbeskjeder angående lokasjonsoppdatering for en mobilnode må autentiseres for å beskytte mobilnoden slik at andre noder på Internett ikke skal kunne ta over mobilnodens data.

4. Linken som mobilnoden som regel er koblet til er ofte trådløs, har mindre båndbredde, og større bitfeil enn faste tilknytninger. Mobilnodene er også ofte batteridrevet (bærbare PC'er), og minimalisering av strømforbruk er også viktig. Derfor bør trafikkbelastningen i båndbredde og tid forårsaket av Mobil IP være så liten som mulig.

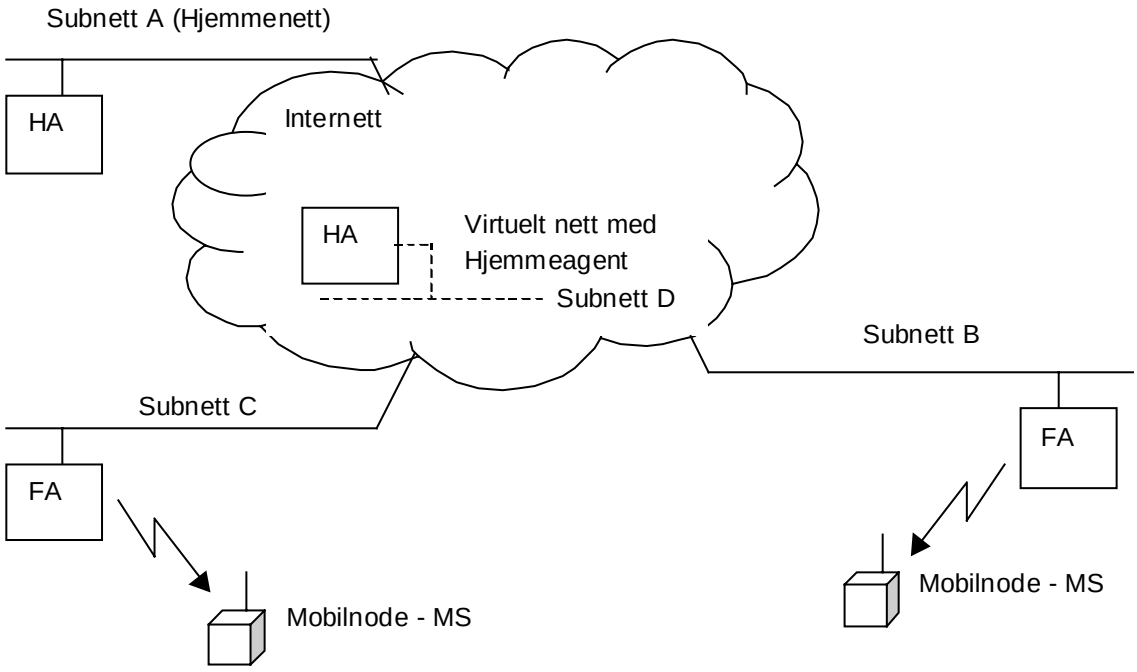
5. Mobil IP må ikke innføre noen restriksjoner når det gjelder tildeling av IP adresser. Det vil si at mobilnoden kan ha en fast IP adresse (slik som noder har i dagens Internett), og at det ikke er nødvendig at mobilnoden benytter en bestemt IP serie for mobilt bruk.

Mobil IP er beregnet for å muliggjøre bevegelse mellom forskjellige subnett, og fungerer like godt med mobilitet mellom heterogene som homogene nett. Det vil si den støtter bevegelser mellom forskjellige ethernet segmenter, samt overganger fra ethernet til for eksempel trådløse LAN.
Mobil IP stiller ingen krav til på lag 2 (Link-laget) for en mobilnode. Dette betyr at den skal være i stand til å håndtere mobilitet uansett hvilken type fysisk nett som noden er koblet til, så lenge en har kontakt med en eller annen form for nett. Det finnes noen lag-2 protokoller som kan håndtere begrenset mobilitet (for eksempel trådløse LAN), og Mobil IP vil fungere uten problemer med disse, siden disse som regel ikke støtter mobilitet mellom forskjellige subnett.

6.2.1 Terminologi

En **mobilnode** (**MS** – *Mobile Station*), er en mobil node eller en mobil ruter som er i stand til å forandre sitt tilknytningpunkt til Internett. En MS har en permanent IP adresse, som kalles hjemmeadresse (*Home address*) som identifiserer mobilnodens hjemmenett, og som ikke endrer seg med nodens bevegelser i nettet. Når mobilnoden er borte fra hjemmenettet, så får den en midlertidig adresse (**COA** – *Care Of Address*), som identifiserer nodens nåværende tilknytningspunkt til nettet, og vil forandre seg ettersom mobilnoden flytter seg fra et tilknytningspunkt til et annet.
Mobilitetsagenter som ofte befinner seg i IP rutere muliggjør og tilbyr mobilitetsstøtte til de mobile brukerne. En **hjemmeagent** (**HA –** *Home Agent*) er en slags ruter som befinner seg i mobilnodens hjemmenett holder orden på MS'ens nåværende tilknytningspunkt i nettet, og er dens stedfortreder når den er borte fra hjemmenettet. Når mobilnoden er borte fra hjemmenettet, så befinner den seg i et fremmed nett, og en har som regel en slags ruter i dette nettet kalt **fjernagent** (**FA** – *Foreign Agent*), som kan videresende pakker til og fra mobilnoden. Som en fellesbetegnelse blir hjemme- og fjernagenter også kalt mobilitetsagenter. En annen node som kommuniserer med en MS blir kalt korresponderende node (**CH** - Correspondent Host).

Som en illustrasjon på hvordan Mobil IP er bygd opp kan en se på figur 6-2.



Figur 6-2 - Mobil IP Illustrasjon

I figuren finnes to fremmede nett, B og C med fjernagenter, og to hjemmenett A og D med hjemmeagenter, samt mobile noder som er tilkoblet fremmednettene med trådløst samband. Data som blir sendt til mobilnodens hjemmeadresse vil bli videresendt til riktig destinasjon i en tunnel (dvs. pakkene blir innkapslet i en ny IP pakke og sendt til MS'ens COA – for mer om dette se kap. 6.4.3).

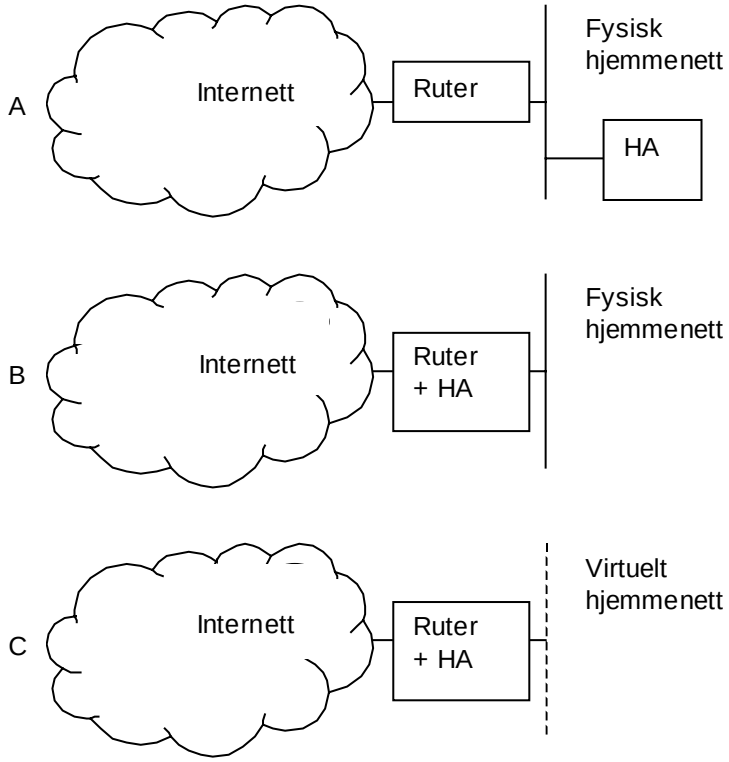
6.3 Mobilitetsagentene

6.3.1 Hjemmeagenten

Hjemmeagentens oppgave er å ta seg av alle pakker til mobilnoden når den ikke befinner seg i hjemmenettet, og videresende disse. For at dette skal være mulig må denne plasseres på en bestemt måte i nettet slik at den er i stand til å være stedfortreder for mobilnoden, noe som kan gjøres på 3 forskjellige måter:

- Plassere hjemmeagenten som en separat node/enhet på hjemmenettet. (A)
- Inkludere hjemmeagenten i ruteren på nettet (B)
- Et virtuelt nett (C)

Dette blir vist i figur 6-3.



Figur 6-3 - Hjemmeagent plassering

Med bruk av virtuelt nett (C) er mobilnoden egentlig aldri tilkoblet hjemmenettet fysisk, og arbeider alltid i fremmede nett. Her er hjemmenettet virtuelt, og benyttes for at en mobilnode som skal kunne nås på en fast IP-adresse. Det virtuelle nettet har ikke et fysisk nett bak ruteren, kun et sett med adresser, og en hjemmeagent. Sett utenfra vil det virtuelle nettet se ut som et fysisk nett.

6.3.2 Fjernagenter

Fjernagenten må plasseres på samme måte som hjemmeagenten i de fremmede nettene som vist i figur 6.3.1-1, unntatt plassering C som ikke eksisterer for fjernagenter da det ikke er mulig for mobile noder å besøke virtuelle nett. Grunnen til at også fjernagenten må plasseres slik er at agenten må kunne nå mobilnoden uten å bruke standard Internett ruting mekanismer. Dette gjelder i situasjoner hvor det er IP-adresse på et av fjernagentens grensesnitt som er mobilnodens COA adresse, og dermed destinasjonen for tunnelen fra hjemmeagenten. I denne situasjonen må fjernagenten videresende pakken direkte til mobilnoden til nodens linklags adresse i stedet. Grunnen til dette er at mobilnoden fortsatt benytter sin faste hjemme IP adresse, og kan derfor ikke nås med standard IP rutingmekanismer, da den benytter en IP adresse som ikke tilhører besøksnettet.
Hvis mobilnoden har fått tildelt en midlertidig IP-adresse (Co-located COA), er den uavhengig av fjernagenten, og kan nås med vanlige IP rutingmekanismer så lenge den befinner seg på det subnettet der COA adressen tilhører.

I Mobil IPv4 vil alle pakker som mobilnoden sender ut inneholde mobilnodens faste hjemmeadresse som avsender adresse, og for at denne metoden skal kunne brukes er en avhengig av at ruteren som tar i mot dataene ikke benytter seg av ingress filtrering. Rutere som bruker ingress filtrering sjekker all pakker som blir mottatt av ruteren, ved å se på avsender IP adressen og hvor den kommer fra. Hvis adressen ikke tilhører det eller de subnettene som finnes innenfor ruteren vil pakken forkastes. Sammen med Mobil IP vil dette

fungere dårlig ettersom pakker fra mobile noder har IP adresser som ikke stemmer med lokale noder, og derfor blir filtrert bort. For å kunne bøte på dette kan en benytte noe som kalles *reverse tunneling* [28], som går ut på at mobilnoden sender alle pakker til fjernagenten, som igjen sender pakkene i tunnel til hjemmeagenten. Hjemmeagenten vil så videresende pakkene til den egentlige destinasjonen. Ytelsesmessig er denne løsningen kanskje ikke den beste da pakker må bevege seg en omvei gjennom nettet for å nå frem. I denne situasjonen kan en bedre løsning være at mobilnoden får tildelt en midlertidig IP adresse (Co-located COA) i stedet.

6.4 Mobil IP virkemåte

Funksjonaliteten til Mobil IP kan ses på som 4 hovedmekanismer:

1. Den må ha en mekanisme som oppdager mobilitetsagenter, kunne finne ut om mobilnoden har beveget seg til et nytt nett, og forstå om den befinner seg i hjemmenettet eller i fremmede nett.

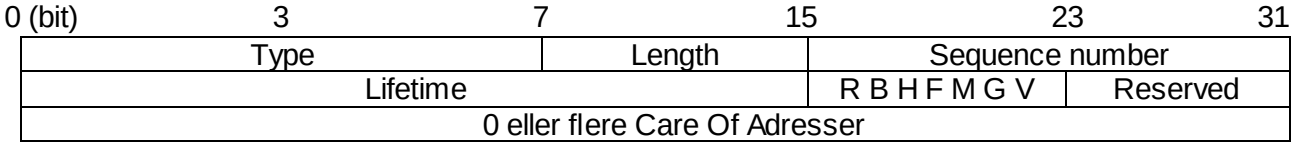
2. Når noden oppdager at den har kommet til ett nytt nett som ikke er hjemmenettet må den skaffe seg en COA adresse.

3. Når mobilnoden har fått en ny COA adresse i det nye nettet må den ha en mekanisme som registrerer den nye adressen hos hjemmeagenten i hjemmenettet, slik at hjemmeagenten vet hvor mobilnoden befinner seg, og kan være dens stedfortreder.

4. Mobil IP må ha mekanismer slik at data sendt til mobilnodens faste adresse blir videresendt og håndtert på riktig måte slik at mobilnåden vil motta dataene på sin nåværende lokasjon.

6.4.1 Agent oppdagelse

I Mobil IP er prosessen for å oppdage mobilitetsagenter omtrent den samme som blir brukt for å oppdage rutere i et nettverk dvs. ved bruk av ICMP (Internet Control Message Protocol) *Router Advertisement* meldinger. For å oppdage agenter benyttes Agent kunngjøringsbeskjeder (*Agent advertisement*) som er modifiserte ICMP beskjeder og blir kringkastet jevnlig ut på nettet for å fortelle om agentens eksistens til mobilnodene.
For å bruke ICMP beskjedene er det lagt til en mobility agent extension i tillegg til den vanlige router advertisement beskjeden: Dette tillegget kan ses i figur 6-4.:



Figur 6-4 – Agent kunngjøringsbeskjed format

Type feltet forteller mobilnoden forteller hva slags tillegg som er lagt til ICMP beskjeden, i dette tilfellet har den verdi 3 som forteller at det er en mobilitetsagent.
Lengde feltet forteller lengden på extension delen.

Sekvensnummerfeltet inneholder et tall som blir økt med en for hver advertisement som blir sendt ut, og brukes for at mobilnodene skal kunne se om agenten har krasjet og restartet, slik at en må registrere seg på nytt.
Lifetime feltet forteller hvor lenge agenten er villig til å tilby tjenesten til mobilnoder før de må registrere seg på nytt (målt i ant. sek.).
Flag feltet inneholder forskjellige bit som kan bli satt:

R – Registrering nødvendig. Registrering med denne fjernagenten er nødvendig, i stedet for å benytte en co-located IP adresse.
B – Busy. Forteller at fjernagenten er opptatt, og ikke er i stand til å ta i mot noen nye noder. (Men den må fortsatt sende ut kunngjøringer på nettet om dens eksistens selv om den ikke har kapasitet til å ta i mot flere, slik at de mobilnodene som allerede er tilkoblet ikke tror at agenten har krasjet, og flytter seg over til en annen agent unødvendig.
H – Hjemmeagent. Kunngjøring som forteller at denne agenten er en hjemmeagent.
F – Fjernagent. Kunngjøring at denne agenten er en fjernagent.
M, **G**, og **V** bitene forteller hva slags innkapsling agenten støtter for pakker som den tar i mot og videresender.

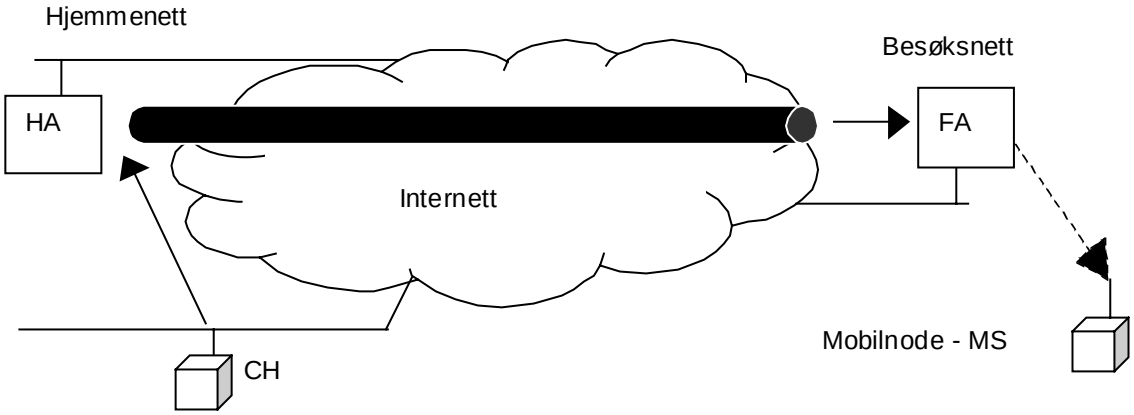
Mobilnoder kan også velge selv å kringkaste en anmodning for oppnå kontakt med en mobilitetsagent. Dette gjøres ved å sende ut en spesiell agent anmodningsbeskjed (*Agent solicitation* – som egentlig er en modifisert ICMP *router solicitation* melding). En mobilnode vil gjøre dette hvis den ikke har mottatt noen agent kunngjøringer innen rimelig tid, for eksempel hvis agenten har krasjet, eller når mobilnoden har byttet nett, og ikke hørt noe fra en agent ennå. Agentene i nettet som mottar denne anmodningen fra en mobilnode vil så svare tilbake til nodens unicast adresse (dvs. svaret er rettet kun til mobilnoden, og ikke til hele nettet).

6.4.2 IP adresse i nytt subnett

Når mobilnoden oppdager at den er kommet til et nytt nett (som ikke er hjemmenettet) må den skaffe seg en ny COA adresse. Hvordan dette skal gjøres blir angitt i agent kunngjøringsbeskjedene (nærmere bestemt R-bitet i *agent advertisement* meldingen).
Har finnes to muligheter:

- MS'en må registrere seg hos en fjernagent og COA blir en IP adresse hos et av agentens grensesnitt.
- Noden må bruke *Co-located* COA, og må skaffe seg en midlertidig IP adresse i det nettet den besøker.

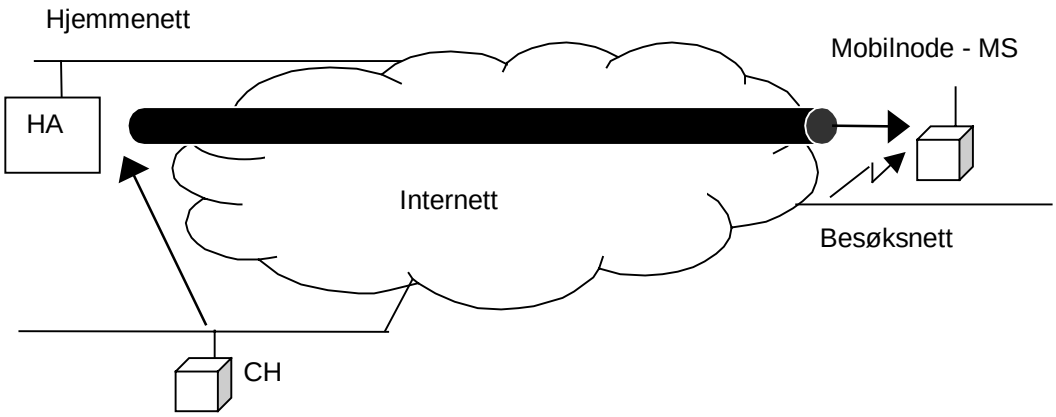
1. I det første tilfellet vil mobilnoden benytte en IP-adresse på et av fjernagentens grensesnitt som COA. Denne adressen blir så oppdatert hos hjemmeagenten, og alle pakker som blir sendt til MS'ens hjemmeadresse vil så bli tatt hånd om av hjemmeagenten, sendt i tunnel til fjernagenten, pakket ut og til slutt sendt til mobilnoden. Dette blir vist i figur 6-5.



Figur 6-5 – Mobil IP pakkeflyt med FA som COA

Fordelen med denne metoden er at mange mobile noder kan dele en COA IP-adresse, og forbruket av IP-adresser blir derfor mindre, noe som er viktig i dagens Internett (IPv4) med det begrensede adresserommet som er tilgjengelig.

2. Det andre tilfellet skaffer mobilnoden seg en midlertidig IP-adresse, såkalt *Co-located* COA, for eksempel ved bruk av DHCP (*Dynamic Host Configuration Protocol*) [33,34]. DHCP er en tjeneste som gir mobilnoden (eller hvilken som helst andre noder i nettet) den nødvendige informasjon for konfigurering av IP adresse. DHCP informasjon blir gitt til mobilnoden av en dedikert server som tildeler IP adresser og andre nødvendige parametre. En DHCP server inneholder en database med alle nødvendige data, og holder streng kontroll over adressetildelingen slik at alle adresser i nettet er unike. Med en *Co-located* COA vil pakker mottatt av hjemmeagenten sendt i tunnel direkte til mobilnodens midlertidige IP adresse, hvor så pakken blir pakket ut av tunnelen. I denne situasjonen vil ikke pakken gå gjennom noen fjernagent, og det er egentlig ikke noe behov for denne i det hele tatt. Dette er vist i figur 6-6:



Figur 6-6 – Mobil IP pakkeflyt med midlertidig (co-located) IP som COA

Ulempen med denne metoden er at mobilnoden opptar to IP-adresser, noe som ikke er gunstig med det begrensede adresserommet som IPv4 har. En fordel med dette er at for oppkoblinger som er kortvarige kan en bruke roamingadressen direkte uten å involvere hjemmeagenten.

6.4.3 Registreringsprosedyre

En hjemmeagent vet mobilnodens nåværende lokasjon ut i fra at mobilnoden sender en registrering til hjemmeagenten når den forandrer lokasjon. Registrering må derfor gjøres hver gang MS'en oppdager en forandring når det gjelder tilkoblingspunkt, hvis fjernagentetn har restartet (dette kan ses ut fra sekvensnummeret i agentkunngjøringene), eller hvis levetiden for den nåværende registreringen snart er utløpt (alle registreringer har en tidsbegrenset levetid, noe som er annonsert i agent kunngjøringene).

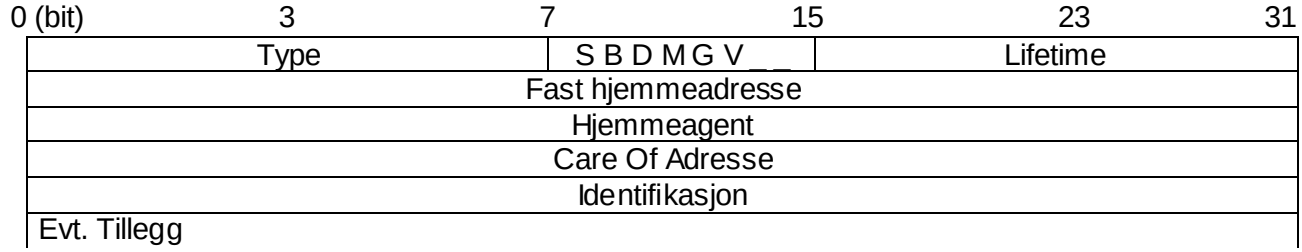
Hvis mobilnoden oppdager ut i fra agent kunngjøringen at den befinner seg i hjemmenettet igjen, vil den avregistrere seg hos hjemmeagenten. HA'en vil da slette alle sine bindinger for den aktuelle MS'en. Etter dette vil mobilnoden oppføre seg som hvilken som helst stasjonær IP node, og den vil benytte seg av ARP, RARP, eller andre mekanismer for å gjøre seg kjent på hjemmenettet igjen.
Hvis mobilnoden finner ut at den befinner seg i et fremmed nett vil den informere sin hjemmeagent om sin nye tilkobling ved å sende en såkalt *registration request* (kalles ofte også binding update). Når HA'en mottar denne vil den svare på registreringen og sende et *registration reply* (også kalt ack) tilbake. Blir *request* beskjeden godkjent vil hjemmeagenten bli mobilnodens stedfortreder i hjemmenettet, og vil videresende alle pakker sendt til nodens faste hjemmeadresse til nodens COA adresse.

6.4.3.1 Registreringsbeskjed format

Registreringer hos hjemmeagenten i Mobil IPv4 skjer via UDP på port 434, og det er definert to typer kontrollbeskjeder:

Type
1 Registreringsforespørsel (*Registration request*)
3 Registreringssvar (*Registration reply*)

Beskjedformatet for en registreringsforespørsel er vist i figur 6-7.



Figur 6-7 – Registration Request format

Innholdet i registreringsbeskjeden er angitt med type 1, og forteller hjemmeagenten om mobilnodens midlertidige adresse (COA), adressen til hjemmeagenten, og den faste hjemmeadressen. I tillegg spesifiseres ønsket levetid for registreringen, bit som settes for å spesifisere hvilke innkapslingsmetoder som kan brukes, om kringkast (*broadcast*) i hjemmenettet skal videresendes til mobilnoden, og om mobilnoden har en midlertidig IP adresse eller ikke (brukes for at hjemmeagenten skal vite hvordan broadcast og multicast skal leveres til mobilnoden).
Beskjedformatet for et registreringssvar er vist i figur 6-8.

0 (bit)	3	7	15	23	31
Type		Code	Lifetime		
Fast hjemmeadresse					
Hjemmeagent					
Identifikasjon					
Evt. Tillegg					

Figur 6-8 – Registration Reply format

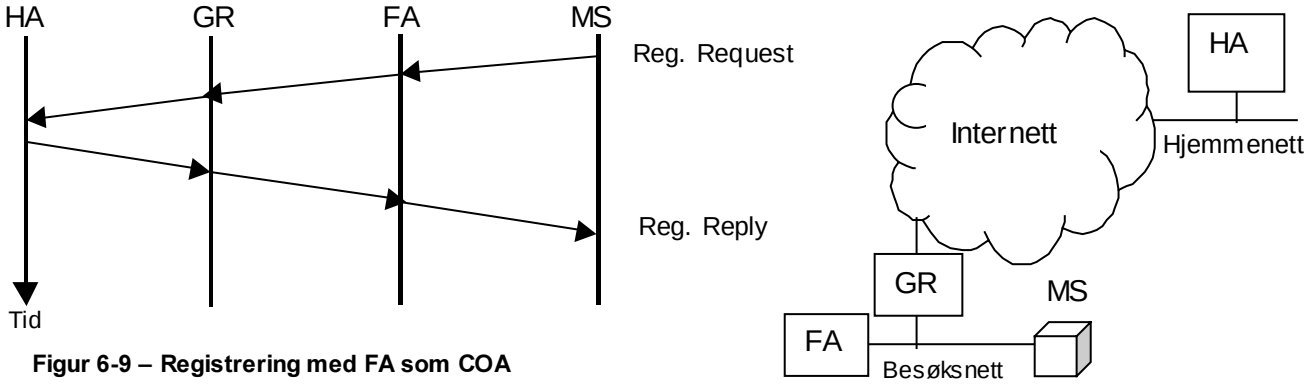
Svaret fra hjemmeagenten er angitt med type 3, og inneholder mobilnodens faste hjemmeadresse, hjemmeagentadresse, identifikasjon, levetid for tjenesten, og et code felt som forteller om registreringen er akseptert eller ikke. Hvis den ikke er akseptert vil nummeret i codefeltet også fortelle hvorfor.
For begge disse pakkene har en også muligheter for tillegg for autentisering av mobilnoden slik at ingen andre skal kunne gi seg ut for å være mobilnoden og ta over dens trafikk (*authentication extension*).

Selve registreringsforløpet med hjemmeagenten kan skje på to forskjellige måter, ettersom en har to muligheter når det gjelder COA adressen (kap. 6.4.2).

6.4.3.2 Registrering ved bruk av fjernagent

Når det er fjernagenten som er mobilnodens COA, vil mobilnoden bruke sin faste IP adresse i alle pakker. For å registrere seg hos hjemmeagenten, må mobilnoden sende en *registration request* til hjemmeagenten, og fortelle hjemmeagenten den nye COA adressen.

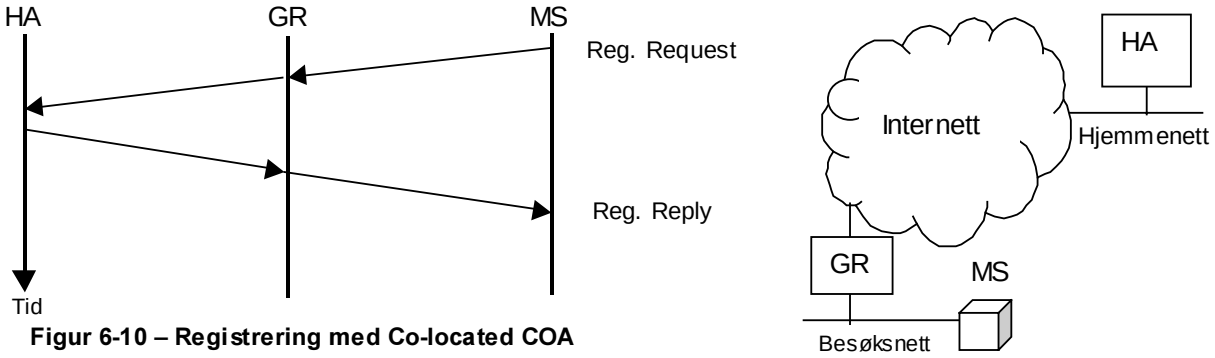
I dette tilfellet sender mobilnoden først registreringsforsepørselen til fjernagenten, som igjen videresender den til hjemmeagenten. Når hjemmeagenten sender svar tilbake vil det bli sendt direkte til fjernagenten (ettersom det er dens IP adresse som er COA), og så videresendt til mobilnoden. Forløpet blir vist i figur 6-9.
Figur 6-9 viser at pakken først blir sendt til fjernagenten (FA), og så til hjemmeagenten (gjennom ruteren GR – *Gateway/Router*). Hjemmeagenten svarer til mobilnodens COA adresse (fjernagenten), og når fjernagenten mottar denne vil den videresende denne til mobilnoden.



Figur 6-9 – Registrering med FA som COA

6.4.3.3 Midlertidig IP adresse

I det andre tilfellet hvor mobilnoden har fått en midlertidig IP adresse (*Co-located Care Of Address*) for bruk i subnettet den besøker. Denne metoden er den enkleste, men gir størst forbruk av IP adresser, men gjør at det ikke lenger er behov for noen fjernagent. Registreringsforespørselen vil her bli sendt direkte til hjemmeagenten i en pakke hvor avsenderadressen er brukerens COA adresse. Figur 6-10 viser forløpet:

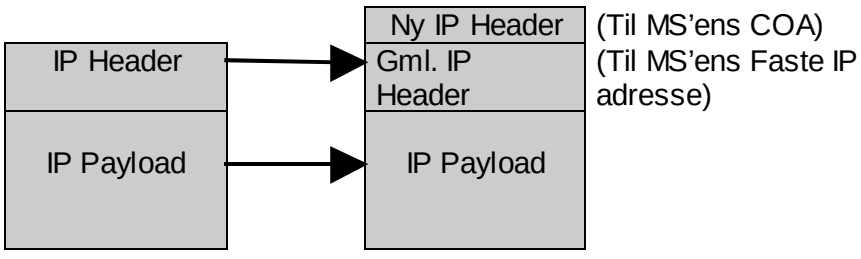


Figur 6-10 – Registrering med Co-located COA

Figuren 6-10 viser hvordan *registration request* beskjeden blir sendt via ruteren til hjemmeagenten (HA). HA'en svarer så med et registreringssvar. I dette tilfellet er det ingen forskjeller når det gjelder trafikkflyten i forhold til annen trafikk mellom stasjonære noder.

6.4.4 Pakkelevering

Hvis MS'en befinner seg i hjemmenettet, vil den sende og motta IP pakker som hvilken som helst stasjonær node. Men når MS'en besøker et fremmed nett, vil alle IP pakker sendt til nodens hjemmeadresse bli tatt i mot av hjemmeagenten (HA). Disse blir da videresendt ved å benytte tunnelering, som betyr at IP pakkene blir pakket inn i nye IP pakker. Destinasjonsadressen i den ytterste IP headeren er mobilnodens COA, og er derfor enden for tunnelen. Når pakken blir mottatt i av destinasjonen vil den ytterste headeren bli fjernet, og videresendt direkte til mobilnoden hvis det er FA'en som er COA, eller begge headerne fjernet og sendt opp til høyere lags protokoller hvis det benyttes co-located COA. Denne formen for innkapsling er vist i figur 6-11.

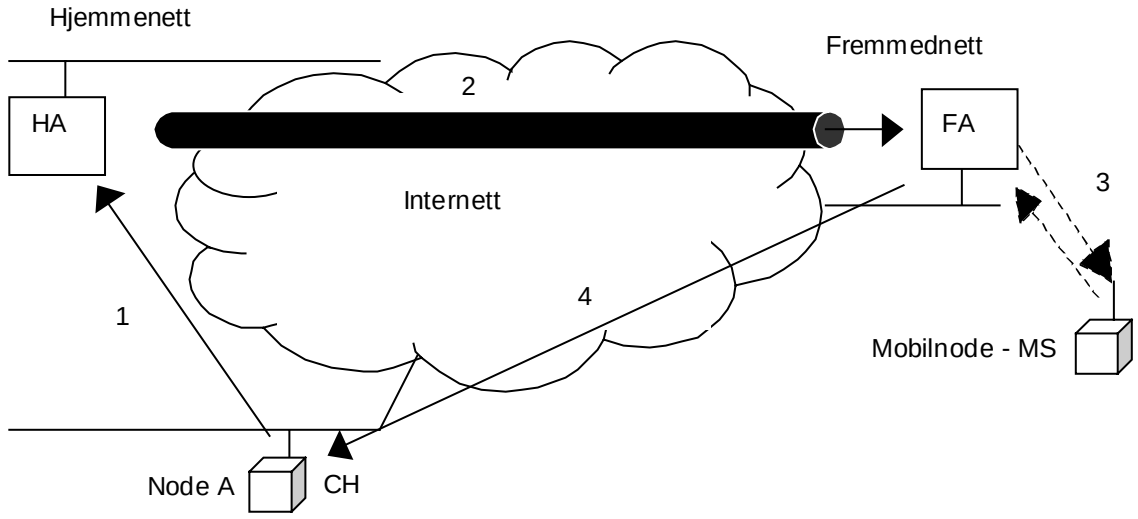


Figur 6-11 – IP i IP innkapsling fra HA til MS

Pakker som blir sendt fra mobilnoden til korresponderende noder blir sendt direkte ved bruk av vanlig IP ruting, unntatt i situasjoner hvor *reverse tunneling* blir benyttet.

6.4.5 Eksempel på Trafikkflyt

I eksempelet under (i figur 6-12) kan en se pakkeflyten for en mobilnode som besøker et fremmed nett, har registrert seg hos hjemmeagenten, og hvor det er fjernagenten som er COA adressen (og dermed enden for tunnelen). Eksempelet viser hvordan en korresponderende node plassert et annet sted (node A – CH) får kontakt med mobilnoden. Flyten av pakker er nummerert i hendelsesrekkefølgen.



Figur 6-12 – Mobil IP pakkeflyt

1. Node A sender en pakke til mobilnodens faste adresse, og denne blir tatt i mot av hjemmeagenten. (Dette skjer via standard IP ruting, ingenting spesielt her).
2. Pakken blir så pakket inn i en ny IP pakke, og sendt i tunnel til COA adressen som i dette tilfellet er fjernagenten.
3. Fjernagenten pakker ut pakken, og leverer den direkte til mobilnoden.
4. Mobilnoden svarer tilbake til Node A via standard IP ruting mekanismer. I dette tilfellet er fjernagenten plassert sammen med ruteren, og pakken går derfor gjennom denne (fjernagenten er her plassert på samme måte som hjemmagenten i figur 6-3 B).

6.4.6 Ruting optimalisering

I eksempelet i kapittelet over vil alle pakker fra den korresponderende noder gå om hjemmeagenten. Denne situasjonen kalles triangulær ruting, og kan være lite gunstig ytelsesmessig da pakkene må gå omveien gjennom hjemmeagenten for å nå frem til mobilnoden. I noen tilfeller kan denne omveien bli veldig lang, for eksempel hvis den korresponderende noden og mobilnoden befinner seg nærme hverandre, mens hjemmeagenten befinner seg langt unna.
For å bøte på dette kan mobilnoden benytte seg av såkalt rutingoptimalisering [27], slik at korresponderende noder slipper å sende all data gjennom hjemmeagenten. Dette gjøres ved at mobilnoden sender en såkalt binding update til alle korresponderende noder. Når den korresponderende noden mottar binding update får den vite mobilnodens COA, og kan sende data direkte til denne adressen i stedet. For å kunne benytte ruting optimalisering må den korresponderende noden være i stand til å benytte seg av denne informasjonen (dvs ha støtte for Mobil IP), noe som dagens enheter ikke har. En trenger derfor en ny eller oppdatert TCP/IP stack i operativsystemet for å kunne håndtere dette. For enheter som ikke støtter dette må pakkeflyt med triangulær ruting gjennom hjemmeagenten brukes.

6.5 Mobil IPv6 kontra Mobil IPv4

Ved utviklingen av Mobil IPv6 brukte en erfaringen fra utviklingen av Mobil IPv4 for å kunne få til ytterlige forbedringer. Mobil IPv6 har mange punkter felles med Mobil IPv4, men i Mobil IPv6 er protokollen fullt integrert med IP(v6), og har en del forbedringer i forhold til Mobil IPv4 [9].

6.5.1 Agent oppdagelse

Oppdagelse av nett fungerer i IPv6 på akkurat samme måte som i Mobil IPv4, ved at mobilnoden lytter etter ”agent *advertisements*”, og ved bruk av ”*agent solicitation*” meldinger om den ønsker det.

6.5.2 IP adresse i nytt subnett

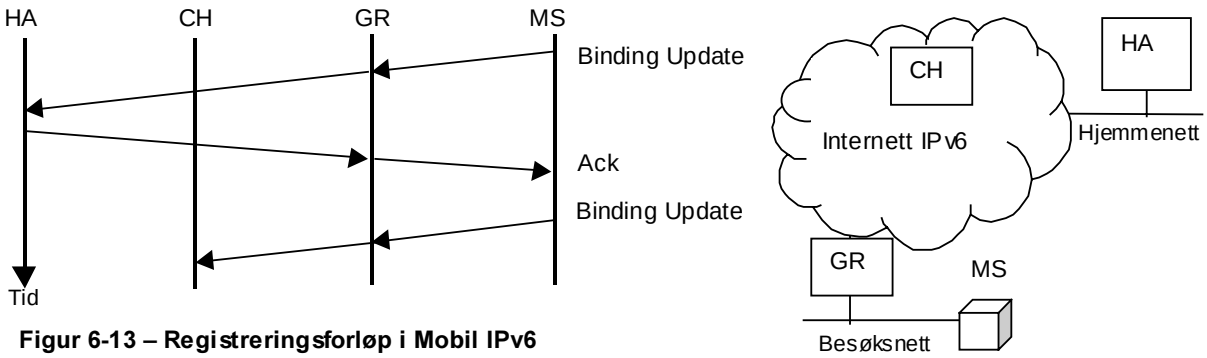
I Mobil IPv6 har ikke lenger behov for spesielle rutere for bruk i fremmede nett dvs. fjernagenter. Virkemåten i IPv6 er på samme måte som når man benytter Co-located COA i Mobil IPv4. For å få tak i ny IP adresse finnes det i Mobil IPv6 to muligheter:

- Mobilnoden får tak i en ny IP adresse gjennom bruk av DHCPv6 (såkalt ”*stateful configuration*” og fungerer på samme måte som DHCP i IPv4).

- Den andre metoden er å benytte ”*stateless autoconfiguration*”. Med *stateless autoconfiguration* genererer mobilnoden sin egen adresse selv uten manuell konfigurering, ekstra servere, og minimal (eller ingen) konfigurering av rutere. Adressen blir generert ut i fra to elementer: lokal tilgjengelig informasjon (fra noden selv), samt informasjon mottatt fra rutere (*Router advertisements*). Den lokale informasjonen er data som identifiserer grensesnittet på noden (Host ID), og er vanligvis hardware (MAC) adressen til nettverkskortet. Den andre delen er ruter delen som identifiserer hvilket subnett noden befinner seg i (Network ID). Disse to elementene så satt sammen til en gyldig IP adresse som kan brukes globalt. Denne metoden kan brukes i nett hvor en ikke trenger noen streng kontroll over hvilke IP adresser som er i bruk, og aksepterer bruk av hvilke som helst adresser så lenge de er unike og rutbare.

6.5.3 Registreringsfasen og ruting optimalisering

I IPv6 er mobilitetsfunksjonaliteten i mye større grad integrert i protokollen. Registreringen foregår på samme måte som i Mobil IPv4 med bruk av Co-located COA. I IPv6 er også rutingoptimalisering innebygd i protokollen, og ikke et valgbart tillegg som i IPv4. Alle IPv6 noder støtter denne funksjonen, og en slipper derfor ulempen med triangulær ruting som i Mobil IPv4 (en annen fordel er at det fører også til mindre last hos hjemmeagenten). Hver gang en mobilnode registrerer seg hos hjemmeagenten, vil den også sende binding updates til korresponderende noder. Forløpet blir vist i figur 6-13.



Figur 6-13 – Registreringsforløp i Mobil IPv6

I IPv6 er ikke lenger kontrollbeskjedene for Mobil IP sendt i egne UDP pakker som i IPv4, men foregår ved bruk av tilleggsheadere (destination headeren) i IPv6. Derfor kan all kontrolltrafikk i Mobil IPv6 legges til i hvilken som helst IPv6 pakke.

6.5.4 Andre forskjeller

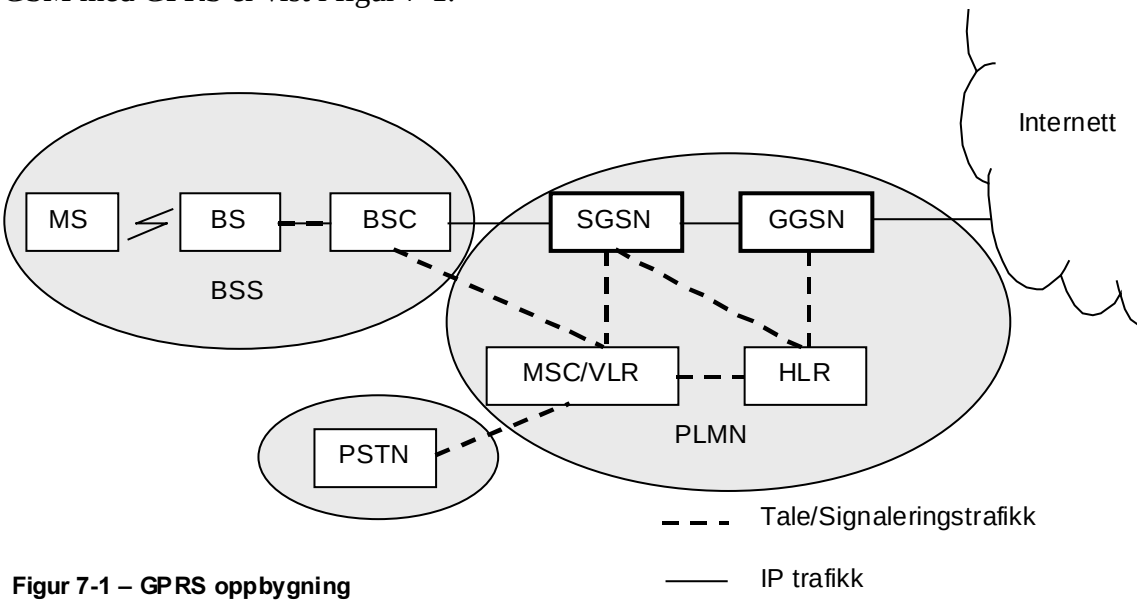
I IPv6 har en tatt hensyn til at mobilnoder som befinner seg i fremmede nett hvor ruterne bruker ingress filtrering. I IPv6 bruker mobilnoden sin midlertidige adresse (COA) som avsenderadresse i IP headeren, og alle pakker vil passere gjennom ruteren uten noen problemer. Den faste hjemmeadressen til mobilnoden blir lagt til i destinasjonstilleggsheaderen i et Home address felt. Dette gjør at den midlertidige adressen er transparent (usynlig) for høyere lags protokoller, ettersom alle IPv6 noder må støtte og kunne prosessere homeaddress feltet i tilleggsheaderen. I IPv4 må en løse dette problemet ved bruk av såkalt *reverse tunneling*, som krever at pakken fra mobilnoden må gå omveien gjennom hjemmeagenten for å nå destinasjonen.
For sikkerhet brukes det i Mobil IPv6 IPsec som er integrert i IPv6, og som har alle nødvendige sikkerhetsfunksjoner (autentisering og kryptering). Mobil IPv4 har en autentiseringsfunksjon som alle Mobil IP noder må støtte, og benytter seg av en MD5 algoritme[1] for dette. Andre funksjoner som kryptering er ikke spesifisert i standarden.
Det er også en rekke andre større og mindre forskjeller på Mobil IPv6 og Mobil IPv4, men de har liten betydning for denne hovedoppgaven og er for øvrig nærmere beskrevet i [9].

7 GPRS & UMTS

7.1 GPRS

General Packet Radio Service (GPRS) [21,22] er en av de nye tjenestene som blir introdusert i GSM Fase 2+ standarden, og vil være tilgjengelig hos enkelte leverandører i slutten av 1999. Hensikten med GPRS er for å støtte periodevis og burst trafikk som en vanligvis har ved pakkeoverføring. GPRS deler og benytter seg av de samme GSM frekvensressursene og TDMA-strukturen (Time Division Multiple Access) som tale og linjesvitsjet datatrafikk benytter.
For å kunne oppnå høyere ytelse blir det i GPRS benyttet en metode som på en dynamisk og fleksibel måte deler de tilgjengelige ressursene med andre GSM tjenester. For å kunne overføre data ved høyere hastigheter blir det brukt flere tidsluker i stedet for en tidsluke per mobilstasjon som dagens GSM gjør. Antall tidsluker som blir benyttet blir regulert dynamisk ut i fra trafikkmengde, ledig kapasitet, og hva slags tjeneste brukeren ønsker og er villig til å betale for. Med GPRS kan en oppnå hastigheter opp mot 115kb/s, i motsetning til dagens datatjeneste i GSM hvor en kun har linjesvitsjet dataoverføring med ytelse på 9,6 kbit/s.

For å tilføre denne funksjonaliteten i det eksisterende GSM nettet, blir det tilført noen nye noder i nettet for å støtte standard protokoller som IP og X.25. GPRS har støtte for både Punkt til Punkt (PTP) og Punkt til Multipunkt (PTM) tjenester, og tilbyr 4 nivåer av tjenestekvalitet (Quality of Service). GPRS er konstruert for å kunne reservere ressurser raskt, og ligger typisk på 0,5 til 1 sekund før noden kan begynne transmisjon av pakker. Oppbygningen av GSM med GPRS er vist i figur 7-1.



Figur 7-1 – GPRS oppbygning

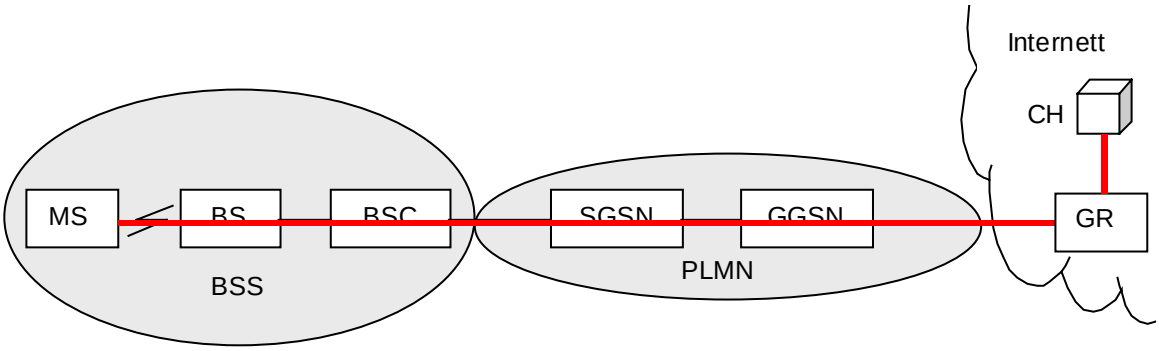
Figuren viser forenklet oppbygning av GSM nettet hvor det fra venstre i figuren begynner med Mobilnoden (MS - Mobile Station), så basestasjonen (BS), og basestasjonskontrolleren (BSC) som er i stand til å kontrollere en eller flere basestasjoner. Alle disse enhetene befinner seg i den delen av nettet som kalles Base Station Subsystem – BSS.
I den landbaserte delen av nettet som kalles *Public Land Mobile Network* (PLMN) befinner de vanlige GSM nettenhetene seg som er: MSC (*Mobile Switching Centre*), VLR (*Visitor Location Register*) og HLR (*Home Location Register*). MSC'en og VLR'et er vanligvis integrert i samme enhet, og har som oppgave å styre en eller flere BSC'er, holde orden på

MS'ene som befinner seg under disse, håndtere signalering, tale og linjesvitsjet datatrafikk. Det er også MSC'en som er koblingen mellom mobilnettet og det vanlige telefonnettet (markert med PSTN – *Public Switched Telephone Network*). HLR'et (Home Location Register) er en slags database som inneholder abonnentdata, og vet til enhver tid hvilket MSC/VLR mobilbrukeren besøker. Hvis en MS får en innkommende samtale fra det faste telefonnettet, blir HLR'et spurt hvor den mobile brukeren befinner seg, og den vil så returnere MS'ens roamingnummer (dvs. hvilket VLR den befinner seg på) tilbake til sentralen som spurte slik at samtalen kan kobles opp.

For å innføre GPRS i GSM nettet blir det lagt til to nye noder: Serving GPRS Support Node (SGSN), og Gateway GPRS Support Node (GGSN). SGSN noden befinner seg på samme nivå som MSC'ene og kommuniserer og holder styr på de enkelte MS'enes lokasjon, og utfører sikkerhetsfunksjoner og tilgangskontroll for GPRS tjenester. GGSN noden fungerer som en gateway/ruter og muliggjør kommunikasjon med eksterne pakkebaserte nett som for eksempel Internett. GGSN noden er koblet sammen med SGSN nodene via et IP basert backbone nett i det landbaserte mobilnettet (PLMN).
HLR'et (Home Location Register) er også utvidet for at den også skal inneholde GPRS abonnentdata. I tillegg kan også MSC/VLR'et utvides for mer effektiv koordinering av GPRS og ikke GPRS tjenester.

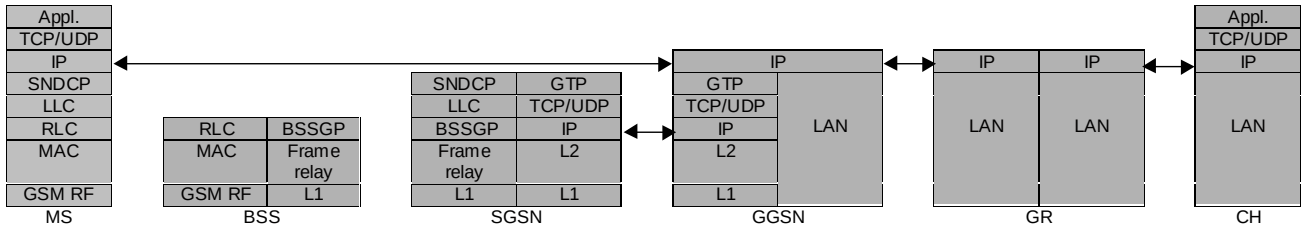
7.1.1 Trafikkeksempel

Hvis en mobilstasjon ønsker å kommunisere med en node som befinner seg på et eksternt nett som Internett, vil trafikkflyten foregå som vist i figur 7-2. Figuren viser den samme GSM nettoppbygningen som i figur 7-1 bortsett fra at det er kun de nodene som overfører IP trafikk som er vist. Figuren viser også en mellomliggende ruter i Internett som ligger mellom gatewayen på GSM/GPRS nettet (GGSN) og den korresponderende noden (CH). I en virkelig situasjon vil det sannsynligvis være flere mellomliggende rutere, men dette har ingen betydning i denne situasjonen. Linjen på figuren markerer IP datatrafikken mellom de to kommuniserende endesystemene.



Figur 7-2 – GPRS og IP trafikk

For den samme situasjonen viser figur 7-3 protokollstakken for MS'en, alle GPRS/GSM noder som har med dataoverføringen å gjøre, en mellomliggende ruter, og den korresponderende noden.



Figur 7-3 – GPRS Protokollstakk

Helt til høyre i figur 7-3 finnes den korresponderende noden (CH) som befinner et sted på Internett. Data fra denne vil passere ruteren (GR), som igjen har direkte kontakt med GatewayGSN noden som er mobilnettets tilknytning til det eksterne nettet (Internett). All IP trafikk til og fra Internett vil i dette tilfellet gå gjennom denne GGSN noden.
Internt i GSM/GPRS nettet vil all IP trafikk mellom GSN noder i nettet benytte GTP (*GPRS Tunneling Protocol*) protokollen. GTP er en tunneleringsprotokoll som går over TCP/IP på det underliggende IP baserte backbone nettet i mobilnettet, og kan ses på figuren der hvor kommunikasjon mellom GGSN noden og SGSN noden som MS'en bruker blir vist. Fra ServingGSN noden, ut på BSS og ut til MS'en vil dataene gå over lavere lags GSM/GPRS spesifikke protokoller, og ute hos mobilnoden er en tilbake på IP nivå igjen.

I situasjoner hvor to GPRS MS'er som befinner seg i samme mobilnett kommuniserer med hverandre vil kommunikasjonen foregå på stort sett samme måten. Forskjellen her er at dataene vil ikke gå innom noen eksterne nett, og vil bare gå over GTP mellom GSN noder internt i mobilnettet. Trafikken vil gå over GTP mellom MS'enes SGSN noder, og gjennom evt. mellomliggende GSN noder trafikken må passere for å komme frem. Fra SGSN nodene og ut til mobilstasjonen vil trafikken foregå på samme måte som vist i figur 7-3.

7.2 UMTS

UMTS (*Universal Mobile Telecommunications System*) er et tredje generasjons mobilkommunikasjons system som er under utvikling. Lanseringen av UMTS er beregnet til å skje i 2002.
Visjonen i UMTS er at den skal bli en fremtidig global plattform for mobilitet med en rekke nye tjenester og forbedringer i forhold til dagens 2. generasjons mobile systemer.
Disse områdene er:

- Forbedrede tjenester når det gjelder:
  - Talekvalitet, radiodekning, og kostnadseffektivitet
  - Tjenestekvalitet (QoS), transmisjonskvalitet, forsinkelse
  - Forbedret effektivitet, og kapasitet (Frekvensallokering og forbruk).
- Nye tjenester og funksjonaliteter:
  - Høy båndbredde for multimedia
  - Båndbredde etter behov, dynamisk allokering av båndbredde, asymmetrisk dataoverføring.
- Fleksibilitet med muligheter for å benytte flere aksessmetoder, frekvensområder, og teknologier.
- Sameksistens, samspill og kompatibilitet med dagens 2. generasjonssystemer, slik at overgangen fra dagens systemer kan foregå flytende.

Arbeidet og utviklingen av UMTS er faseinndelt, og er i dag kun fase 1 som til en viss grad er klarlagt. I fase 1 er det spesifisert funksjonalitet for høykvalitetstale ved bruk av lave bitrater, dataoverføring med hastigheter fra 144kbit/s og opptil 2Mbit/s som skal være godt egnet for multimedia, roaming mellom GSM og UMTS nettverk, og Dual mode/band GSM/UMTS telefoner for bruk i begge nettverk.
I motsetning til da GSM ble utviklet er man i utviklingen av UMTS klar over behovet for datatjenester, men hvordan dette skal gjøres er ikke spesifisert ennå, men at UMTS vil kunne tilby tjenester godt egnet for IP trafikk er meget sannsynlig.
Hvordan nettene skal være bygd opp er heller ikke bestemt ennå, bortsett fra at de kan være basert på eksisterende nett for eksempel GSM, ISDN og Internett.

8 Trådløse LAN - IEEE 802.11 & HIPERLAN

8.1 Trådløse LAN

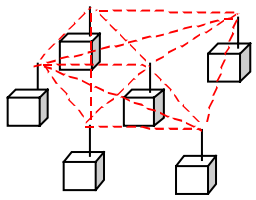
Trådløs kommunikasjon er en teknologi hvor det er og har vært stor utvikling de siste årene. Denne teknologien har gitt oss muligheten for nettverkstilgang uten å være fysisk tilkoblet til et nett. Trådløse lokalnett (Wireless LAN), er som deres trådbaserte motparter (LAN) utviklet for å gi brukerne høy båndbredde i et begrenset geografisk område. Et WLAN har vanligvis en dekning på typisk 100m innendørs, avhengig av bygningsmasse, struktur osv. Denne typen nett er godt egnet i miljøer hvor en ikke har eksisterende kabling for trådbaserte nett (noe som er dyrt og vanskelig å legge opp i ettertid), midlertidige nett, samt i miljøer hvor en har behov for mobilitet.
På dette området er det i to dominerende teknologier: IEEE 802.11 Wireless LAN [16,17,18], og det europeiske systemet HIPERLAN [16,17]. Disse to er standarder som opererer på lag 1 (fysisk) og lag 2 (link laget) dvs. under nettverkslaget (IP).

8.1.1 Wireless LAN - IEEE 802.11

IEEE (*Institute of Electrical and Electronics Engineers*) begynte utarbeiding av standarden 802.11 [16,17,18] i 1990, og ble ferdig standardisert i 1997. 802.11 tilbyr en hastighet på 1Mbit/s og med muligheter for utvidede hastigheter på 2Mbit/s og 11Mbit/s. Standarden er i stand til å benytte seg av tre typer fysiske lag:

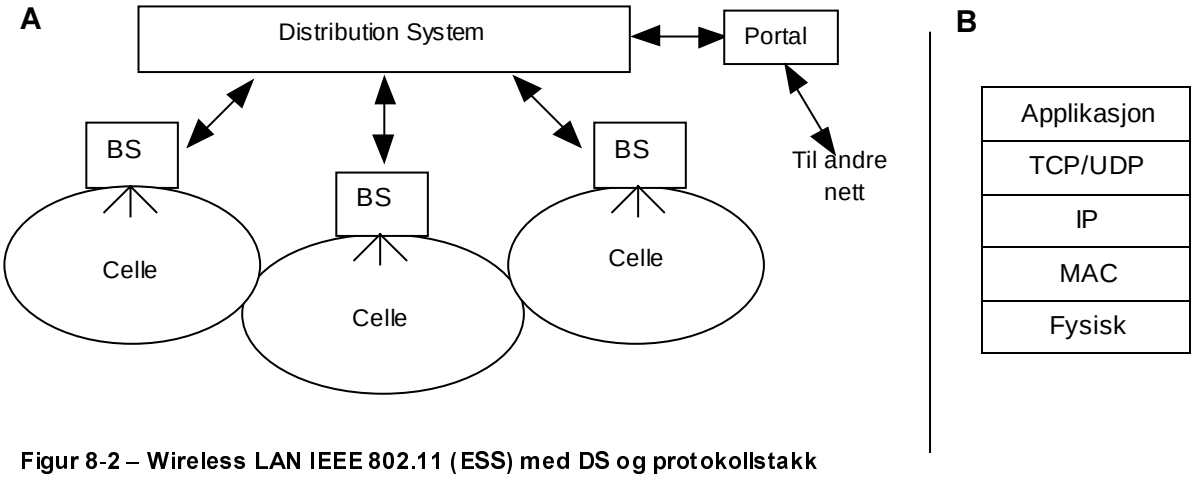
- DFIR (Diffuse Infra-Red) er basert på infrarød kommunikasjon, og er beregnet for bruk i kontorlandskap, og tilbyr en rekkevidde på opptil 10m.
- DSSS (*Direct Sequence Spread Spectrum*) er radiobasert, benytter seg av det frie 2.4GHz båndet, og har maksimum rekkevidde på noen hundre meter med fri sikt, eller flere kilometer ved bruk av direktive antenner.
- FHSS (*Frequency Hopping Spread Spectrum*) er også radiobasert, og benytter seg av samme frekvensområde som DSSS, og har tilsvarende rekkevidde.

Forskjellen mellom de to radiobaserte lagene er at FHSS kan være noe bedre i miljøer hvor en har mange celler som overlapper hverandre, fordi frekvenshopping vil føre til mindre interferens fra nabosendere, og at de to metodene håndterer forskjellige typer støy på forskjellige måter.
I 802.11 standarden kan det benyttes to forskjellige topologier: en hvor en gruppe av stasjoner kommuniserer direkte med hverandre i et ad-hoc nettverk (kalles BSS – *Basic Service Set*). Her må alle nodene høre hverandre for å kommunisere og en er helt uavhengig av noen infrastruktur, noe kan være nyttig i sammenhenger hvor en trenger midlertidige nettverk f.eks. i konferanserom, messer osv. Et slikt nett er vist i figur 8-1.



Figur 8-1 – IEEE 802.11 ad-hoc nettverk (BSS)

Den andre topologien som kan ses i figur 8-2A har mobilnodene tilgang til et backbone nett via tilgangspunkter (access points) - dvs. basestasjoner (denne topologien kalles ESS – *Extended Service Set*). Denne typen kan brukes til å gi kontinuerlig netttilgang ved å utplassere flere basestasjoner med radiodekning som overlapper hverandre slik at en får kontinuerlig dekning i et bestemt geografisk område. For å kunne få til dette har 802.11 et håndteringssystem for brukerne i de forskjellige cellene. Dette blir realisert med noe som kalles distribusjonssystem (DS) som har to hovedoppgaver: Den må vite nodens nåværende posisjon innenfor en gruppe av celler som DS'en har ansvar for (dvs. bygge opp svitsjetabeller, samt sørge for styring av handover når nodene flytter seg fra celle til celle). Den andre oppgaven er å transportere pakker til riktige celler (dette gjelder både for intern trafikk mellom celler, og ekstern trafikk til og fra det faste nettet.



Figur 8-2 – Wireless LAN IEEE 802.11 (ESS) med DS og protokollstakk

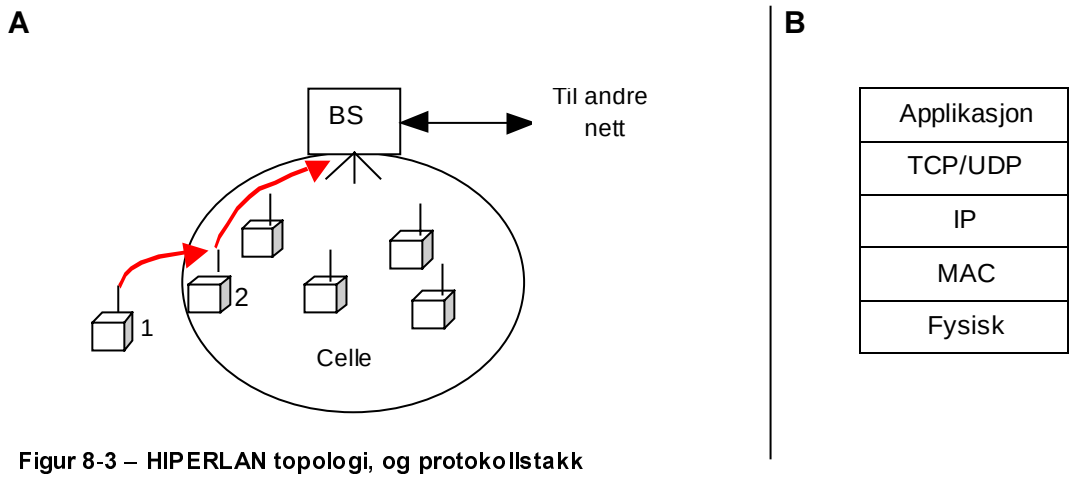
Aksessmetoden i 802.11 er i MAC-laget (*Medium Access Control*) som befinner seg på lag 2 (figur 8-2B) og benytter seg av såkalt CSMA/CA (*Carrier Sense Multiple Access with Collision Avoidance*) for aksesskontrollen til mediet. I CSMA/CA vil noder som har data å sende lytte på radiokanalen (som er felles for alle noder) og høre om det er noe trafikk. Hvis kanalen er ledig, vil nodene som ønsker å sende data vente en tilfeldig valgt periode, og så sende dataene hvis kanalen fortsatt er ledig. Dette blir brukt for å redusere antall kollisjoner slik at ikke alle noder som har noe å sende begynner å sende samtidig. Et problem med CSMA/CA metoden er at det kan oppstå problemer med skjulte noder. Situasjonen oppstår hvis to noder er utenfor hverandres radiorekkevidde, men innenfor rekkevidden til basestasjonen. Hvis de to nodene vil sende data samtidig vil begge tro at kanalen er ledig ettersom de ikke hører hverandre, mens hos basestasjonen vil det oppstå kollisjoner. Det mottatte signalet hos basestasjonen vil da være en blanding av de to nodenes radiosignaler, som blir helt uforståelig og forkastes. For å bøte på dette er det lagt til en tilleggsmekanisme som kan brukes når det er mye kollisjoner. Denne mekanismen går ut på at nodene som vil sende sender en forespørsel om å få lov til å sende (RTS – *Request To Send*), og får da svar når den får lov til å sende (CTS – *Clear To Send*). Når CTS meldingen blir sendt ut til en node, vil også alle andre noder innenfor radiorekkevidde motta denne meldingen, og vente med å sende.
Adresseringen i 802.11 er litt forskjellig fra vanlig MAC adressering. Vanlig MAC adressering benytter to felter – avsender og mottaker, og siden denne adresseringen ikke er i stand til å indikere en mobilnodes posisjon i en gruppe av celler, er det lagt til tilleggsinformasjon for å få til dette. 802.11 standarden har derfor i stedet 4 adresse felter, som har forskjellige meninger avhengig av pakkens retning.

Handover i wireless LAN kan løses på forskjellige måter, og er ikke nærmere spesifisert i standarden. Løsningene her blir derfor proprietære, dvs avhengig av hver enkelt leverandør,

og en har ingen garanti for at en kan bygge opp et nett med celler bestående av komponenter fra forskjellige produsenter. Handover og ruting mellom celler i 802.11 foregår kun på lag 1 og 2, og er derfor uavhengig av IP. For å kunne ha kontinuerlig IP tilknytning uten Mobil IP må alle celler tilhøre det samme IP subnettet. Ettersom handover mellom 802.11 celler vil foregå på samme IP subnett er det ikke behov for handover på IP nivå med Mobil IP. Handoverytelsen i 802.11 kan variere ettersom standarden ikke spesifiserer hvordan handover skal gjøres. Dette blir derfor opp til hver enkelt produsenten å bestemme hvordan dette blir implementert, som gjør noen produkter kan gi ha mindre avbrudd og pakketap enn andre ved handover.

8.1.2 HIPERLAN

HIPERLAN (HIgh PErformance Radio LAN) [16,17] er en ETSI-standard (*European Telecommunications Standards Institute*) for trådløse LAN hvor målet var å gi ytelse som var på linje med trådbaserte nett som Ethernet. Standarden spesifiserer et radiobasert trådløst LAN som benytter frekvenser i 5 GHz området, tilbyr en datahastighet på 23.529 Mbit/s, og har en rekkevidde på opptil 50 m. HIPERLAN har også støtte for multihopp ruting, og tidsbundne (multimedia) tjenester ved å kunne sette prioritering på kanaltilgangen. Det finnes ingen form for garantert tjenestekvalitet, kun en ”*best effort*” tjeneste, og hvis mange noder sender såkalte høyprioritetspakker vil man få de samme problemene som hvilket som helst annet delt medium.
Aksessmetoden på MAC-laget (figur 8-3B) i HIPERLAN er basert på en ”*carrier sense*” mekanisme som fungerer på stort sett samme måten som i 802.11 (CSMA/CA). Den fungerer slik at når en node ønsker å sende data vil den lytte på kanalen, og hvis kanalen har vært fri for en fast bestemt periode kan den begynne å sende. Hvis kanalen ikke er fri når den begynner å lytte eller at en annen node begynner å sende først, vil den vente til kanalen blir ledig, og så gå over til å benytte en annen aksessmetode for å kunne sende. I denne aksessmetoden finnes det 3 faser som mobilnoden må igjennom før den kan sende: en prioritetsfase, en elimineringsfase, og en vikepliktsfase. I prioritetsfasen vil de nodene som har satt høyest prioritet på pakkene vinne. De to siste fasene sørger for at det kun er én node som kommer til å sende data. Med denne aksessmetoden er sannsynligheten for kollisjoner meget liten (under 3%).
Også i HIPERLAN har en mulighet for de to typene topologier som i 802.11: ad-hoc nettverk uten basastasjon og direkte kontakt mellom noder, samt nettverk med basestasjon og tilgang til backbone nett. I HIPERLAN finnes det ingen mulighet for at flere celler kan fungere sammen ved bruk av et distribusjonssystem for handover og svitsjing av pakker slik som i 802.11. I stedet finnes det en mulighet for at noder kan videresende (*forwarding*) data for andre noder for å øke rekkevidden. Figur 8-3A viser oppbygningen av et HIPERLAN med basestasjon og mobilnoder som både befinner seg utenfor og innenfor rekkevidde av basestasjonen.



Figur 8-3 – HIPERLAN topologi, og protokollstakk

I HIPERLAN finnes det to forskjellige typer noder: Forwarder og ikke-forwarder. Ikke-forwardere vet bare om andre stasjoner den har direkte kontakt med (dvs. stasjoner som er innenfor radiorekkevidde), mens de som er forwardere vet om hele nettverkstopologien. Informasjon om topologien blir mottatt og vedlikeholdt kontinuerlig ved at kontrollpakker blir sendt regelmessig mellom nodene. Hvis en node som befinner seg utenfor rekkevidde av basestasjonen eller en annen node som den vil snakke med må den benytte forwarding for at pakken skal komme frem. Pakken må da sendes gjennom en forwarder som noden har kontakt med, eller sende en kringkasting til alle nabostasjonene. Hver pakke blir så videresendt fra forwarder til forwarder, enten via såkalt *unicast relaying* (videresending til bestemte noder), eller via kringkasting til den når destinasjonen eller pakkens levetid løper ut. Figur 8-3A viser et slikt tilfelle hvor noden som er markert med 1 befinner seg utenfor rekkevidde av basestasjonen, og node 2 er en forwarder som videresender dataene for node 1 (videresendingen av data er markert med piler).
Et problem med denne metoden er at en baserer seg på at andre stasjoner er villige til å donere energi og prosesseringskraft til nytte for andre stasjoner. Det kan være vanskelig å finne "frivillige" som vil ha den rollen i et miljø hvor ressursene mange ganger er begrenset. Ressursene det er snakk om her er tilgjengelig båndbredde og energi. Derfor kan det skje at noder ikke vil være forwarder slik at så få noder som mulig deler den tilgjengelige båndbredden. Når det gjelder energi så vil en node som er forwarder og videresender mye pakker for andre noder bruke mer energi, noe som er lite ønskelig da det som regel er snakk om utstyr som er batteridrevet.
En fordel med denne løsningen er at en ikke er avhengig å installere ny infrastruktur for å gi økt dekning, da rekkevidden kan økes med bruk av forwarding, noe som ikke er mulig i 802.11.

9 Mobil IP - Handover

Handover i Mobil IP vil oppstå hver gang en mobil node beveger seg fra et IP subnett til et annet IP subnett. Etterhvert som trådløs lokalnettilgang blir mer vanlig vil noder som beveger seg rundt bytte nett forholdsvis ofte (noe som er avhengig av nodens fart og retning). Grunnen til dette er at disse nettene som regel har små celler med rekkevidde som ligger i området noen få meter og opptil noen hundre meter. Hver gang noden forandrer tilknytning over til et annet nett som tilhører et annet IP subnett må det foretas en Mobil IP handover. Denne prosessen tar tid, og kan gi problemer i form av redusert tjenestekvalitet for brukeren.

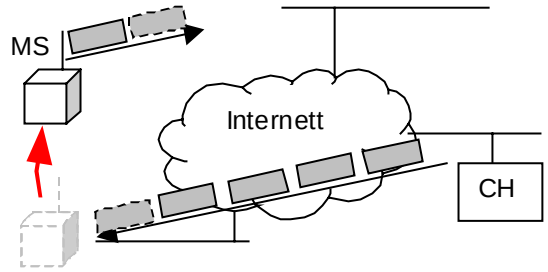
Handoverprosessen i Mobil IP kan deles opp i følgende faser:

- Oppdagelse av nytt nett.
- Oppnå ny COA adresse.
- Registrering hos HA gjennom evt. FA for lokasjonsoppdatering

Hver av disse fasene vil ta en viss tid å gjennomføre, og det en ønsker er at tiden som brukes er så kort som mulig.
Slik som Mobil IP [25] er i dag, blir handover foretatt når noden finner ut at den befinner seg i et nytt nett ved å lytte etter *agent advertisements*. I ICMP *router advertisement* delen av *agent advertisements* meldingene finnes det et levetidsfelt som angir hvor lang levetid hver kunngjøring har. Dette feltet forteller hvor ofte en advertisement blir sendt ut på nettet, og blir benyttet for å oppdage at noden har mistet kontakt med nettet. Hvis noden ikke har mottatt en ny advertisement melding innenfor det tidsrommet som er annonsert i ICMP lifetime feltet vil noden anta at nettet ikke er tilgjengelig lenger.
Hvis mobilnoden mottar en *agent advertisement* fra en annen mobilagent enn den er tilknyttet for øyeblikket vil den gå ut ifra at den har kommet på et nytt nett.
Når noden har mottatt en *agent advertisement* og funnet ut at den er på et nytt nett, vil den få tak i en ny COA adresse. Etter dette er gjort vil den foreta en Mobil IP lokasjonsoppdatering (dvs. sende en binding update) hos hjemmeagenten muligens via en fjernagent. (Hvordan dette skjer er beskrevet i kap. 6.4).

9.1 Problemet med Mobil IP handover

Ettersom den prosessen som er beskrevet ovenfor vil ta en viss tid, kan den føre til problemer for brukeren. Slik handover er spesifisert i Mobil IP i dag vil det føre til et avbrudd i forbindelsen for mobilnoden, og alle pakker som noden sender ut i perioden etter at den har mistet forbindelsen, og før den oppdager at nettet er borte vil gå tapt fordi ingen ruter eller andre lokale noder vil motta pakkene. Også i motsatt retning vil det samme problemet oppstå ved at pakker fra korresponderende noder vil gå tapt da mobilnoden ikke er i stand til å ta i mot pakker på den gamle COA adressen lenger. Denne situasjonen er vist i figur 9-1.



Figur 9-1 – Mobil IP pakketap

For TCP forbindelser [1] vil det ikke bare oppstå problemer akkurat under avbruddet, men pakketapet vil også føre til ettervirkninger. Dette er fordi TCP protokollen vil retransmitere alle pakker som har gått tapt, og på grunn av virkemåten til TCP's mekanismer for hastighetsregulering. For at TCP skal utnytte den tilgjengelige datahastigheten gjennom nettet reguleres datahastigheten på grunnlag av pakketapet. Hvis en pakke går tapt tror protokollen at det er metning (*congestion*) i nettet (stor trafikk), og setter så ned datahastigheten. Etter pakketapet vil hastigheten gradvis økes igjen til et pakketap oppstår igjen. Hvis en mobilnode foretar handover ofte vil ytelsen for TCP forbindelser kunne bli betraktelig redusert, da TCP er slik at den vil bruke litt tid på å komme opp i den hastigheten den var på før avbruddet. Tiden det tar før hastigheten er oppe igjen er avhengig av båndbredden og forsinkelsen over nettet, og kan også variere noe da det finnes forskjellige versjoner av TCP som reagerer litt forskjellig når det oppstår pakketap.
For protokollen UDP vil ikke de samme problemene som i TCP oppstå da den ikke har noen retransmiteringsalgoritme, og blir vanligvis anvendt til andre formål enn TCP. UDP benyttes ofte i såkalte realtime applikasjoner, hvor en er avhengig av liten eller relativ konstant forsinkelse i nettet for at applikasjonen skal kunne fungere tilfredsstillende. Tap av data og forsinkelse forårsaket av Mobil IP kan derfor gi problemer for realtimeapplikasjoner.
En type applikasjoner som faller under denne kategorien er for eksempel toveis tale programmer (kalles også VoIP – Voice over IP) som benytter seg av H.323 protokollen (for eksempel Microsoft Netmeeting). Her blir Internetts pakkebaserte "*best effort*" tjeneste benyttet for overføring av en samtale, i motsetning til det vanlige linjesvitsjede telefonnettet hvor datahastigheten og forsinkelsen hele tiden er konstant. VoIP applikasjoner er veldig følsomme når det gjelder variasjon i forsinkelsen mellom endesystemene, og denne typen tjenester får problemer og blir vanskelig å bruke hvis det oppstår store variasjoner i forsinkelsen. For å kompensere og håndtere dette har H.323 standarden av avanserte *codecer* som komprimerer data slik at det ikke er behov for høye bitrater, samt buffring og intelligent koding slik at en er i stand til å tåle noe tap av pakker og forsinkelser. Hvis forsinkelsene og avbruddene går utenfor de rammene algoritmene er i stand til å håndtere vil det oppstå avbrudd i lyden, og brukeren vil oppleve tjenesten som dårlig eller ubrukelig. En metode som til en viss grad kan hjelpe på dette er å øke buffringen av data, men dette innfører også økt forsinkelse noe som gjør det vanskeligere å føre en samtale.

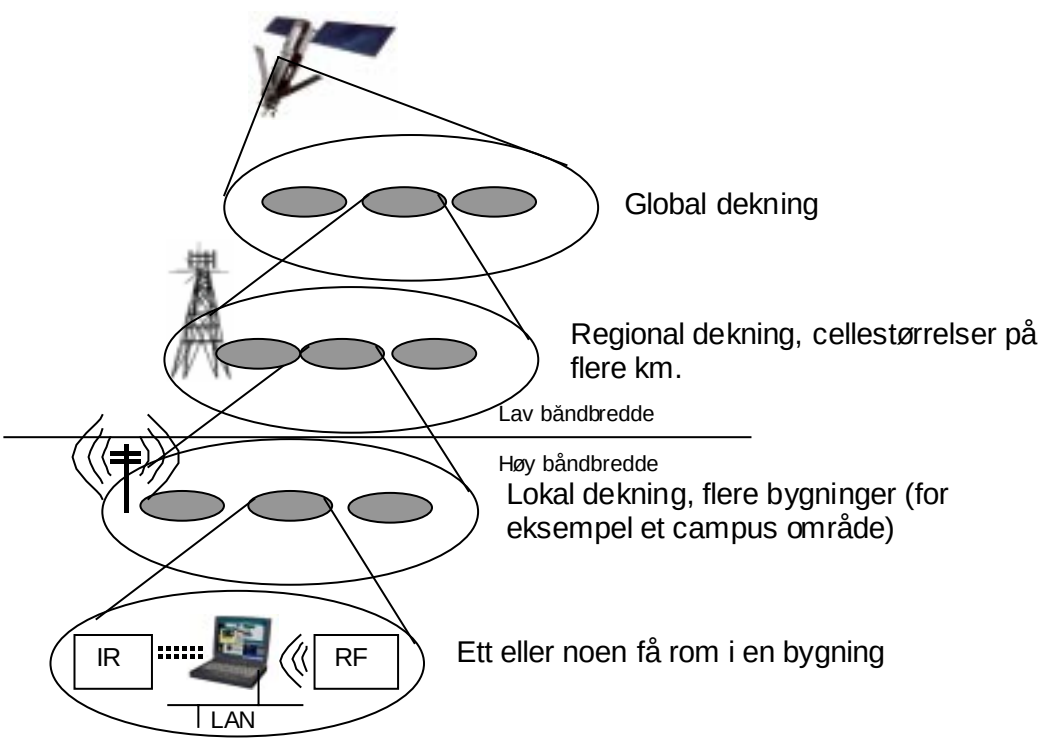
9.2 Teknologier og handover

Hvis Mobil IP skal brukes som en fremtidig plattform for mobilitet og kontinuerlig nettilgang, så må de mobile enhetene være utstyrt med flere nettverksteknologier, da det ikke finnes et system som dekker alle krav. Ønsket er å ha kontinuerlig nettilgang med høy båndbredde hele tiden, men det er dessverre ikke mulig. I dag finnes det ikke ett system som gir kort forsinkelse, høy båndbredde, stort dekningsområde, og kapasitet til mange brukere på en gang, så derfor er man i stedet avhengig av å bruke flere forskjellige teknologier avhengig av brukerens lokasjon. Når mobilnoden befinner seg innenfor et bygg eller campus område som har dekning av trådløse LAN vil disse benyttes, og når noden befinner seg utenfor dekning av

disse lokale nettene må den gå over til vanlige mobilnett og evt. satellitt for å kunne opprettholde nettilgangen. Det finnes et stort spekter av teknologier tilgjengelig i dag, og flere er på vei, hvor båndbredden i systemene varier i fra noen få kbit/s og opptil flere titalls Mbit/s. På tross av denne mangfoldigheten av teknologier kan de stort sett deles i inn to hovedkategorier: Nett som gir høy båndbredde over et lite geografisk båndbredde, og nett som gir lav båndbredde over et stort geografisk område. For at en mobilnode skal kunne ha tilgang til høy ytelse, og mulighet for kontinuerlig nettilgang må den derfor ha grensesnitt for flere typer nett.
I Mobil IP så må dette håndteres på en måte slik at brukeren får en nettverkstilknytning som hele tiden er best mulig. For å kunne håndtere dette er det hensiktsmessig å organisere de forskjellige teknologiene i et slags hierarki [14] basert på deres båndbredde og dekningsområde, og kunne skille mellom forskjellige typer handover.
Motivet for å bruke Mobil IP for handover i denne sammenheng er at Mobil IP er uavhengig av underliggende teknologi, noe som de fleste andre mobilhåndteringssystemer ikke er. Mobil IP er også en åpen standard, og kan benyttes av alle. Ettersom Mobil IP er integrert i IP og er transparent for høyere lag vil den kunne benyttes med de samme høyere lags protokoller og applikasjoner som blir brukt i stasjonære noder. Det finnes ingen andre alternativ i dag som innehar alle disse fordelene.

9.2.1 Hierarki med teknologier

Hierarkiet kan bygges opp slik at en legger de teknologiene som har høy båndbredde og lav rekkevidde nederst, og de nettene som har størst dekningsområde og lavest båndbredde øverst i hierarkiet. Denne oppbygningen blir vist i figur 9-2.



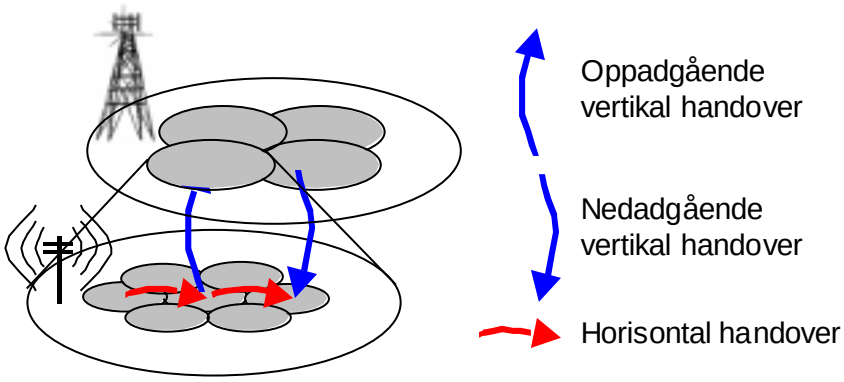
Figur 9-2 – Hierarki med teknologiene ut i fra deres dekningsområde

Helt nederst på figuren finnes teknologier som benytter fast trådbasert tilknytning som for eksempel ethernet som gir en båndbredde på 10 eller 100Mbit/s. De neste er trådløse LAN

teknologier som 802.11 og HIPERLAN og benytter IR eller RF teknologi som gir forholdsvis høy båndbredde (noen Mbit/s) innenfor en rekkevidde på noen ett eller noen få rom i en bygning.
På neste nivå finnes andre trådløse nett som gir høy båndbredde, og med rekkevidde på noen hundre meter utendørs, for eksempel 802.11 med bruk av direktive antenner for å dekke et campus område.
Neste nivå er teknologier som faller inn i den andre kategorien hvor båndbredden er drastisk redusert, men hvor dekningsområdene er mye større. Her finnes konvensjonelle mobile nett som for eksempel GSM med GPRS som kan gi brukeren tilsvarende ytelse som med dagens ISDN (hvordan dette kan kombineres med Mobil IP omtales i kap. 11). På det øverste nivået i hierarkiet finnes satellittbasert teknologi som gir global dekning, men gir meget lav båndbredde og har som regel også høye brukskostnader. Et eksempel på dette er Iridium hvor datahastigheten i første generasjon ligger på 2.4kbit/s.

9.2.2 Horisontale og vertikale handover

Ut i fra inndelingen av de forskjellige teknologiene i et hierarki som beskrevet i forrige kapittel kan handover deles opp i to hovedkategorier: horisontale og vertikale handover. Forskjellen på disse blir vist i figur 9-3.



Figur 9-3 – Horisontale og vertikale handover

Horisontale handover er som figur 9-3 viser handover som skjer når en mobilnode beveger seg mellom celler innenfor en type teknologi. Eksempler på dette er handover når en MS beveger seg mellom to celler i et GSM nett, og handover som skjer når en mobilnode beveger seg mellom celler innad i et 802.11 WLAN. Denne typen handover skjer alltid på lavere lag, og har som regel ingen innvirkning på IP nivå, og krever derfor ikke noen handover på Mobil IP nivå ettersom en vanligvis beveger seg innenfor samme subnett.

Den andre typen handover som kalles vertikale handover skjer når en mobilnode foretar en handover mellom to forskjellige teknologier og beveger seg vertikalt i hierarkiet. Et eksempel på dette kan være at en mobilnode beveger seg ut av dekningen for et WLAN og må gå over til å bruke GSM tilknytning i stedet. Denne typen handover vil også føre til handover på Mobil IP nivå.

Som også vist på figur 9-3 kan vertikale handover også deles opp i to kategorier: oppadgående vertikal handover og nedadgående vertikal handover. Oppadgående vertikale handover skjer når en mobilnode foretar handover til teknologier som ligger ovenfor i hierarkiet som har større dekningsområde og mindre båndbredde. Nedadgående vertikale handover er handover som blir foretatt til nett som ligger under i hierarkiet som har mindre dekningsområde og

høyere båndbredde. Disse to vil skje på forskjellige grunnlag, og blir derfor nærmere beskrevet hver for seg.

9.2.2.1 Oppadgående vertikal handover

Når mobilnoden beveger seg ut av dekningsområdet av det nettet/teknologien den bruker for øyeblikket og oppdager dette vil den gå over til å benytte seg av et nett høyere i hierarkiet for å kunne fortsette å ha nettilknytning. Slik vil mobilnoden fortsette oppover i hierarkiet hver gang den mister dekningen i det nettet den for øyeblikket er tilknyttet. Som eksempel på dette kan en tenke seg en situasjon hvor en mobil bruker beveger seg vekk fra sin kontorplass, ut av bygningen, og ønsker å være oppkoblet på nett hele tiden. På kontoret vil mobilnoden være tilknyttet med fast trådbasert ethernet. Når brukeren kobler noden av nettet vil mobilnoden gå over til bedriftens trådløse LAN som har dekning i hele bygget. Etter hvert beveger brukeren seg ut av bygget og radiodekningen til det trådløse LAN'et, og må derfor gå over til å benytte mobile nett som for eksempel GSM.

9.2.2.2 Nedadgående vertikal handover

For brukeren er det ønskelig å ha en best mulig tjeneste til enhver tid, ikke bare kontinuerlig nettilgang. Ettersom det er de nettene som er lavest i hierarkiet som gir de beste tjenestene, er det nettet som er tilgjengelig lavest i hierarkiet som er foretrukket av brukeren. Derfor må mobilnoden være i stand til å bevege seg nedover i hierarkiet igjen når nett med bedre ytelse blir tilgjengelig. For eksempelet beskrevet i forrige kapittel vil dette kunne være når mobilnoden kommer tilbake på kontoret, og vil gå over fra å bruke GSM til å benytte seg av det trådløse LAN i bygningen som gir mye bedre ytelse (og sannsynligvis lavere brukskostnader).
Her er problemstillingen litt annerledes enn for oppadgående v. handover: Mobilnoden har fortsatt kontakt med et nett høyere oppe i hierarkiet, men har beveget seg inn i dekningsområdet for et lavere nett som gir bedre ytelse, og ønsker derfor å benytte dette i stedet. Den eneste måten å finne ut om et nett lavere i hierarkiet er tilgjengelig på er å kontrollere grensesnittet og sjekke om det er noe nett tilgjengelig. Det beste måten å gjøre dette på ytelsesmessig er å ha alle grensesnitt på mobilnoden påslått til enhver tid, slik at handover kan foretas med en gang nett blir oppdaget på lavere nivå. I praksis blir ikke dette noe særlig gunstig, på grunn av høyt strømforbruk hvis alle grensesnitt er aktive til enhver tid. Da en mobilnode vanligvis er batteridrevet bør derfor bør bare det eller de grensesnittene som er i bruk være påslått.
Da det er sannsynligvis ikke er nødvendig for en mobilnode å komme på nett med samme sekund et lavere nett blir tilgjengelig, kan en metode basert på polling benyttes i stedet, som går ut på at hvert grensesnitt for teknologier lavere i hierarkiet blir slått på med jevne mellomrom for å sjekke om det finnes nett tilgjengelig på en av disse. Hvis mobilnoden poller hvert grensesnitt med intervall på 10 sekunder mellom hver gang, vil strømforbruket fortsatt kunne være relativt lavt, og at mobilnoden kun vil ha en maksimal forsinkelse på 10 sekunder når det gjelder oppdagelsen av et bedre nett, noe som ikke burde være et problem for noen.

9.2.3 Mobil IP og vertikale handover

Slik som Mobil IP er i dag har den ingen mekanismer for å kunne håndtere flere nett slik som beskrevet når det gjelder vertikale handover. I dag vil noden velge det grensesnittet som den hører noe fra først, og fortsetter å bruke det så lenge det er dekning. Derfor må funksjonaliteten beskrevet i kapittel 9.2.1 og 9.2.2 implementeres i Mobil IP, slik at flere grensesnitt (figur 9-4) kan håndteres på en hensiktsmessig måte.

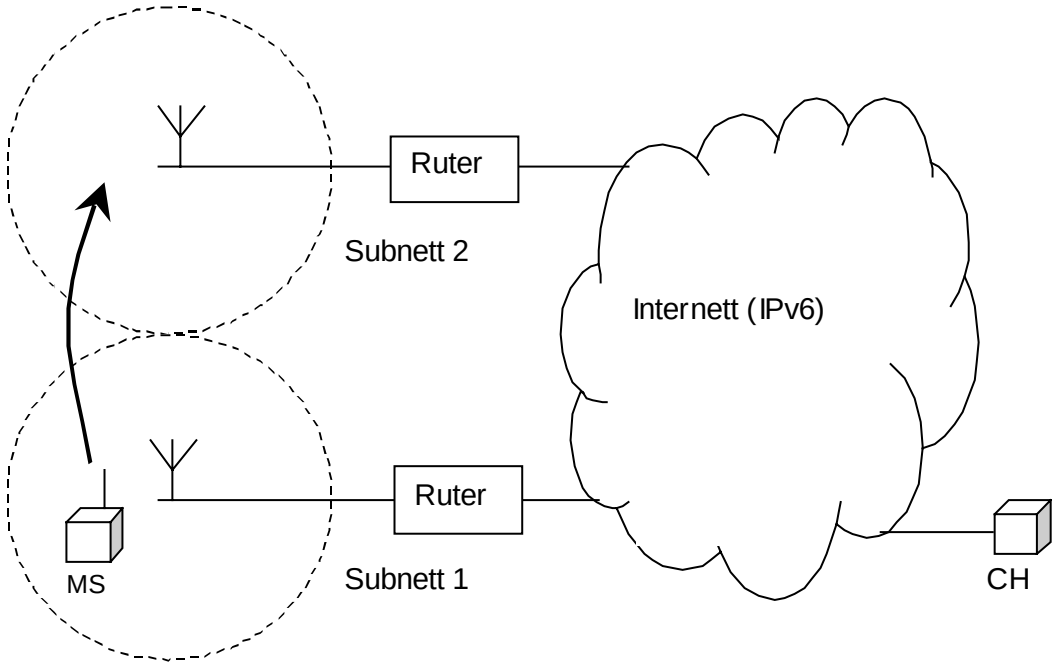
Applikasjon				
TCP		UDP		
IP (inkl. Mobil IP)				
Ethernet	802.11 IR/RF	HIPER-LAN	GSM/ GPRS	Satellitt

Figur 9-4 – IP protokollstakk med flere grensesnitt

9.3 Handover i Mobil IPv6 med bruk av anycast

I IPv6 ble det introdusert en adresseringsform som ikke eksisterer i IPv4. Denne heter anycast [5], og som tidligere fortalt (kap. 5.4.3), gir anycast muligheten til å sende data til en node blant en gruppe av noder som har en felles adresse. Noe av hensikten med oppgaven var å undersøke om bruken av anycast for adressering av en gruppe av rutere kan gjøre handover enklere eller raskere.
For å undersøke dette betraktes to forskjellige subnett (med to forskjellige teknologier), som har hver sin ruter som kan treffes på den samme anycast adressen.

Situasjonenen er at en mobilnode (MS) befinner seg i subnett 1, og kommuniserer over et trådløst LAN og Internett med en annen node (CH).
Situasjonen som oppstår er at mobilnoden flytter seg ut av dekningsområdet for det trådløse LAN'et, og beveger seg over i et annet subnett (som i dette tilfellet bruker en annen type nett – f.eks. HiperLAN). Hvilken type teknologi som brukes i disse to nettene er egentlig likegyldig, bortsett fra at det skal være to forskjellige subnett med to forskjellige rutere, som benytter den samme anycast adressen.
Denne situasjonen er illustrert i figur 9-5.



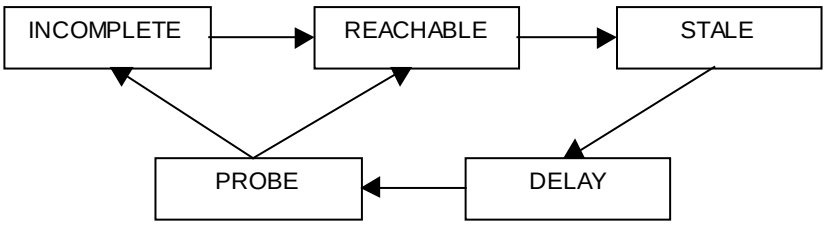
Figur 9-5 – Handover mellom to subnett, og rutere med felles anycast adresse

Når mobilnoden beveger seg over i det nye nettet, har den ikke oppdaget at den er i et nytt nett ennå, og hensikten med å studere denne situasjonen er å se om den nye ruteren i det andre

nettet bare kan ta over pakkene som mobilnoden sender til CH, ettersom ruteren som mobilnoden sender til har den samme IP adressen.
For å kunne finne ut hva som vil skje her må en se på hvordan pakker blir sendt mellom noder lokalt i et nettverk. Med lokalt menes det noder (for eksempel datamaskiner og rutere) som befinner seg på samme fysiske (sub)nettet. Det som skjer i denne situasjonen er at pakker som sendes til lokale noder vil bli sendt direkte til destinasjonene med deres lag 2 adresser. For å finne ut hvilken lag 2 adresse lokalnodene har benyttes det i IPv6 noe som kalles *Neighbour Discovery*.

9.3.1 Neighbour Discovery

Neighbour Discovery i IPv6 [6] har funksjonalitet som integrerer de tre IPv4 funksjonene: ARP (*Address Resolution Protocol*), ICMP Router Discovery beskjeder, og ICMP Redirect beskjeder.
Neighbour Discovery blir benyttet av noder for å finne linklagsadresser for nabonoder, og finne ut om de fortsatt er tilgjengelige.
I dette tilfellet er det *address resolution* og *neighbour unreachability detection* mekanismene i som har betydning. *Address Resolution* (AR) benyttes for å finne assosiasjoner mellom en nodes IP adresse og lag 2 (MAC) adresse, mens *neighbour unreachability detection* (NUD) benyttes for å finne ut at en nabo (en lokal node) ikke lenger er tilgjengelig. Neighbour discovery behandler anycast adresser på samme måte som unicast adresser (med 2 unntak, men disse har ingen betydning i dette tilfellet [6]).
For å finne en assosiasjon mellom en linklagsadresse og en IP adresse vil AR søke gjennom en *neighbour cache*, som er en liste over lokale noders IP adresser og deres linklagsadresse sammen med et statusfelt. Statusfeltet innehar en av 5 mulige tilstander: INCOMPLETE, REACHABLE, STALE, DELAY, og PROBE. Hvordan en post beveger seg mellom disse tilstandene er vist i tilstandsdiagrammet i figur 9-6.



Figur 9-6 – Neighbour cache tilstander

Denne oversikten er meget forenklet, og det finnes også andre muligheter for å bevege seg mellom tilstandene (blant annet timeouts, og ved mottak av advertisements i de forskjellige tilstandene), men figuren viser de som er mest relevant for dette tilfellet.
Det som skjer når IP laget vil sende en pakke til en lokal node vil *Address Resolution* delen av *neighbour discovery* sjekke om den finner en linklagsadresse ved å søke gjennom *neighbour cachen*. Hvis det finnes en post i cachen med en linklagsadresse så kan pakken sendes til destionasjonen med en gang.
Dersom MS beveger seg fra subnett 1 til 2 slik som i figur 9-5, og den ennå ikke har oppdaget at den er i et nytt nett, vil MS'en fortsette å sende pakker til CH via ruterens anycast adresse. Dessverre så vil dette føre til at alle pakker gå tapt ettersom ruteren i subnett 2 har en annen linklagsadresse enn den i subnett 1. Adressen som mobilnoden sender til er den linklags-adressen MS'en fant ut at ruteren hadde når den befant seg i subnett 1, og når mobilnoden kommer over i subnett 2 vil den fortsette å sende til den samme lag 2 adressen (ettersom det er denne adressen som finnes i *neighbour cachen*), og ingen pakker kommer frem.

Her vil *neighbour unreachability detection mekanismen* slå til, og etter en kort stund vil den aktuelle posten i cachen gå over til STALE tilstand ettersom noden ikke lenger mottar noe data fra den aktuelle adressen. Etter en viss tid vil den gå over i DELAY tilstand, som betegner at noden ikke lenger er tilgjengelig. Etter en timeout hvor noden er sikker på at noden ikke er tilgjengelig vil noden gå over til PROBE tilstand, og sender ut en forsepørsel (*neighbour solicitation*) for å finne ut om det eksisterer noen noder lokalt med denne IP adressen. Hvis mobilnoden får svar vil posten gå over til REACHABLE tilstand, og hvis ikke INCOMPLETE. Denne prosessen vil ta såpass lang tid at innen den er i stand til å finne den nye lag 2 adressen til ruteren vil mobilnoden ha oppdaget at den er på et nytt nett.

Ut i fra dette kan en derfor konkludere med at bruken av anycast har ingen betydning i vertikale handover, og bevegelser mellom subnett, ettersom en er avhengig av at IP adressen (uansett om den er uni- eller anycast) må kobles til riktig hardware adresse.

10 Handover – forbedringer

For å kunne foreslå forbedringer for handover i Mobil IP kan det være hensiktsmessig å separere de forskjellige fasene som mobilnoden går igjennom når handover foretas, og se på hvert område hver for seg. Det mest praktiske er da å se på hendelsene hver for seg i den rekkefølgen hver hendelse oppstår.
Hendelsesforløpet er som følger:

Mobilnoden beveger seg ut av dekningsområdet for sin nåværende tilknytning, og må derfor bytte over til et annet nett basert på en annen teknologi. Mobilnoden må derfor gå over til et annet interface, og komme seg på nett igjen via denne.
Etter mobilnoden har fysisk/linklagstilknytning på det nye nettet, må den motta en agentkunngjøring, og skaffe seg en COA adresse slik som angitt i kunngjøringsbeskjeden. Den vil så foreta en lokasjonsoppdatering med Mobil IP ved å sende en binding update til hjemmeagenten og evt. andre noder.

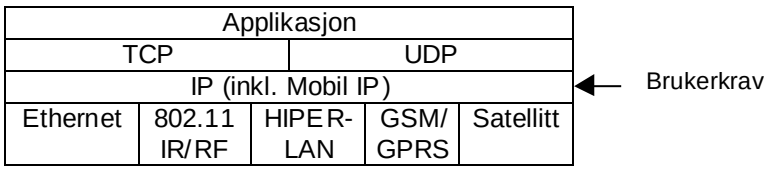
Oppsummert kan dette ses på som to hovedhendelser:

- Vertikal handover
- Mobil IP handover

Målet i denne sammenheng er at handover skal foregå på hurtigst mulig måte, med minst mulig pakketap, minimalisere trafikkbelastningen forårsaket av handover trafikk, og at løsningen bør være skalerbar og kunne fungere i alle mulige typer nett, datahastigheter, og avstander (forsinkelser).

10.1 Forbedringer i vertikale handover

For å kunne håndtere vertikal handover best mulig er det ønskelig med en funksjonalitet enten integrert i eller ved siden av Mobil IP som styrer handoverprosessen på en måte som tilfredsstiller brukerens krav. Ettersom brukerens krav kan variere bør funksjonaliteten i Mobil IP kunne konfigureres av brukeren ut i fra bestemte kriterier som for eksempel handoverytelse, strømforbruk, foretrukket nett og pris (de forskjellige nettene vil sannsynligvis ha forskjellige priser når det gjelder anvendelsen av dem). Oppbygningen av Mobil IP og de forskjellige grensesnittene er vist i figur 10-1.



Figur 10-1 – IP protokollstakk med flere grensesnitt

For å kunne forbedre oppadgående vertikale handover er det ønskelig at Mobil IP også er i stand til å utnytte linklagsinformasjon fra grensesnittene som for eksempel linkstatusdata som signalstyrken. Hvis mobil IP overvåker signalstyrken for grensesnittet som er i bruk, kan den være i stand til å forberede seg på en mulig handover, slik at den slår på grensesnitt som ligger høyere i hierarkiet blir når signalstyrken til det nåværende grensesnittet går under en viss verdi. Hvis så signalstyrken går under enda en grenseverdi hvor ikke signalstyrken er god nok

til å gi en pålitelig tjeneste vil Mobil IP gå over til å benytte det andre grensesnittet. På denne måten kan mobilnoden være i stand til å komme seg på et nytt nett før dekningen forsvinner helt på den nåværende tilknytningen, og dermed få til en såkalt ”soft handover” som er handover helt uten pakketap. For å finne ut mer nøyaktig hvordan dette kan implementeres slik at vertikale handover blir foretatt best mulig kan for eksempel kunnskap og erfaring fra handover i mobiltelefonsystemer benyttes.
I tillegg til å benytte seg av linklagsinformasjon kan Mobil IP benytte seg av høyerelags informasjon som ruter advertisement og agent advertisement meldinger, hvor intervallet mellom hver beskjed er angitt i meldingen. Hvis mobilnoden ikke mottar en kunngjøring når det er forventet bør den slå på grensesnittet over for å være forberedt på en mulig handover. Handover bør ikke foretas med en gang da dette ikke behøver å bety at nettet er på vei til å forsvinne, men at pakken ikke ble mottatt på grunn av en kollisjon.

10.2 Forbedringer i Mobil IP Handover

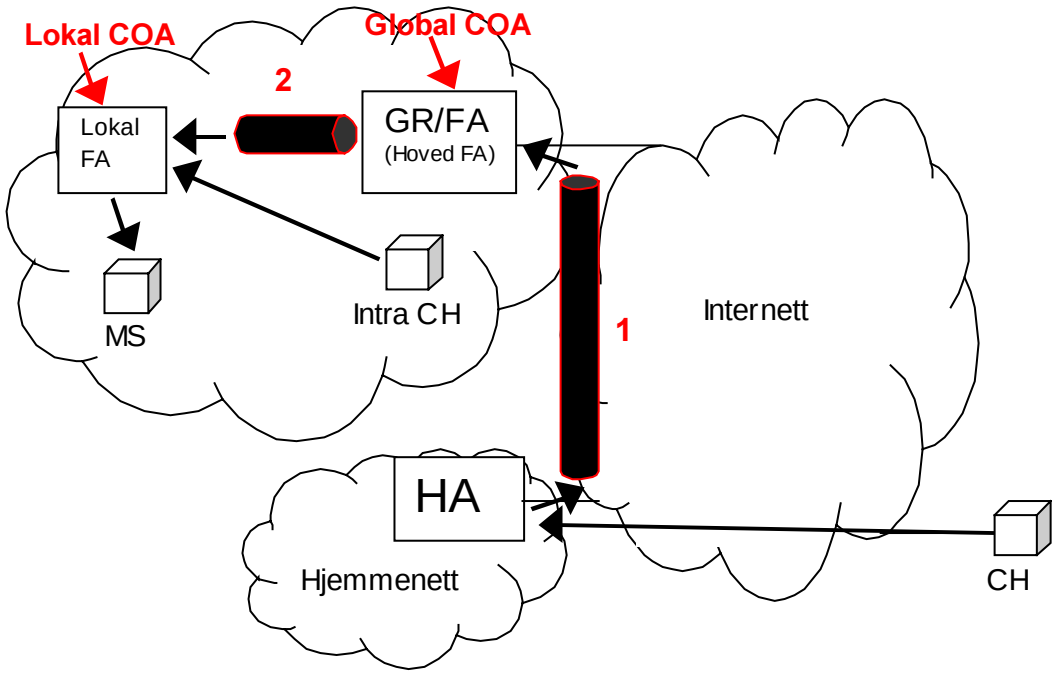
Etter at mobilnoden har utført den vertikale handoveren og kommet seg inn på det nye nettet på linklagsnivå, må noden få tilbake tilknytningen på nettverkslaget, noe som er litt forskjellig i IPv4 og IPv6. Der hvor det er forskjeller er det beskrevet i kap. 10.2.5.

10.2.1 Agent oppdagelse

For å kunne komme seg på nett igjen må mobilnoden lytte etter en *agent advertisement* for å få tak i en ny COA adresse. For å kunne oppdage det nye nettet så fort som mulig er det ønskelig at advertisement beskjedene sendes ut så ofte som mulig, men ikke så ofte at den vil oppta noen større del av kapasiteten på linken [15]. Hvor ofte disse kan sendes ut er avhengig av hastigheten i nettet, og i Mobil IP standarden er det anbefalt at de ikke blir sendt ut oftere enn en gang per sekund.
Når mobilnoden mottar *agent advertisement* meldingen vil den se om den må registrere seg via fjernagenten, og bruke dens IP adresse som COA, eller om den skal bruke co-located COA via DHCP. Da det er ønskelig at dette skal skje hurtigst mulig er det ønskelig å benytte seg av FA’en som COA, da dette vil gå raskere enn å benytte DHCP.
Da det er ønskelig at mobilnoden mottar *agent advertisement* så fort som mulig kan også noden benytte seg av *agent solicitation* meldinger. Dette kan benyttes slik at når mobilnoden må bytte over til et annet grensesnitt sender den ut en *agent solicitation* melding ut på det nye nettet med en gang, ettersom den da vet at den er på et nytt nett.
Fasen etter at mobilnoden har fått seg en COA adresse er den som er kritisk. Her må mobilnoden sende en binding update til hjemmeagenten, noe som kan ta lang tid og er vanskelig å gjøre noe med ettersom mobilnoden her er prisgitt den forsinkelsen over nettet. Hvis avstanden til hjemmeagenten er lang, og trafikken høy (større forsinkelse) vil registreringen kunne ta lang tid.
For å bøte på dette kan noden samtidig som den sender binding update til hjemmeagenten sende en binding update til den forrige FA’en den brukte slik at pakker som i mellomtiden blir sendt til den gamle COA adressen vil bli tatt hånd om og videresendt til den nye COA adressen.
Ettersom Mobil IP slik den er spesifisert i dag håndterer mobilitet likt uansett om det er små bevegelser innenfor en site, eller globale forflytninger. For å gjøre handover raskere kan en løsning være å benytte Mobil IP i et **hierarki** av mobilitetsagenter [12, 15]. En slik løsning vil kunne gjøre handover raskere, gi mindre belastning på nettene ettersom færre handoverregistreringer må sendes til hjemmeagenten, og gjøre Mobil IP mer skalerbar i større sammenhenger.

10.2.2 Hierarki

Ved å innføre et hierarkisk system med mobilitetsagenter [12, 15] vil det mellom mobilnoden og hjemmenettet være flere tunneler og agenter i motsetning til dagens situasjon hvor det kun er én tunnel og én agent. Tunnelene vil ligge etter hverandre, og for situasjoner hvor mobilnoden beveger seg lokalt innenfor et administrativt domene, vil kun den siste tunnelen flyttes. Et administrativt domene vil typisk være en samling av flere nett innenfor en site, for eksempel et universitetskampus eller en større bedrift. En lokal bevegelse vil være usynlig for hjemmeagenten og evt. korresponderende noder utenfor domenet, og disse trenger derfor ingen lokasjonsoppdatering.
Globale forflytninger, dvs. forflytninger mellom forskjellige domener vil situasjonen foregå som før (som i Mobil IP i dag) hvor hjemmeagenten og alle korresponderende noder må oppdateres.
Oppbygningen av et slikt hierarki blir vist i figur 10-2.



Figur 10-2 – Hierarki av tunneler

Figur 10-2 viser et eksempel hvor det er benyttet to fjernagenter i stedet for en, noe som fører til at det er to IP tunneler (lag 3) etter hverandre mellom mobilnoden og hjemmeagenten. I denne situasjonen må mobilnoden også ha to COA adresser:

- En lokal COA i et subnett innenfor siten som er adressen til en lokal FA, eller mobilnodens IP adresse hvis co-located COA benyttes. Det er denne adressen som forandres når mobilnoden flytter seg innenfor siten, og er enden for den andre tunnelen (markert med 2 på figur 10-2).

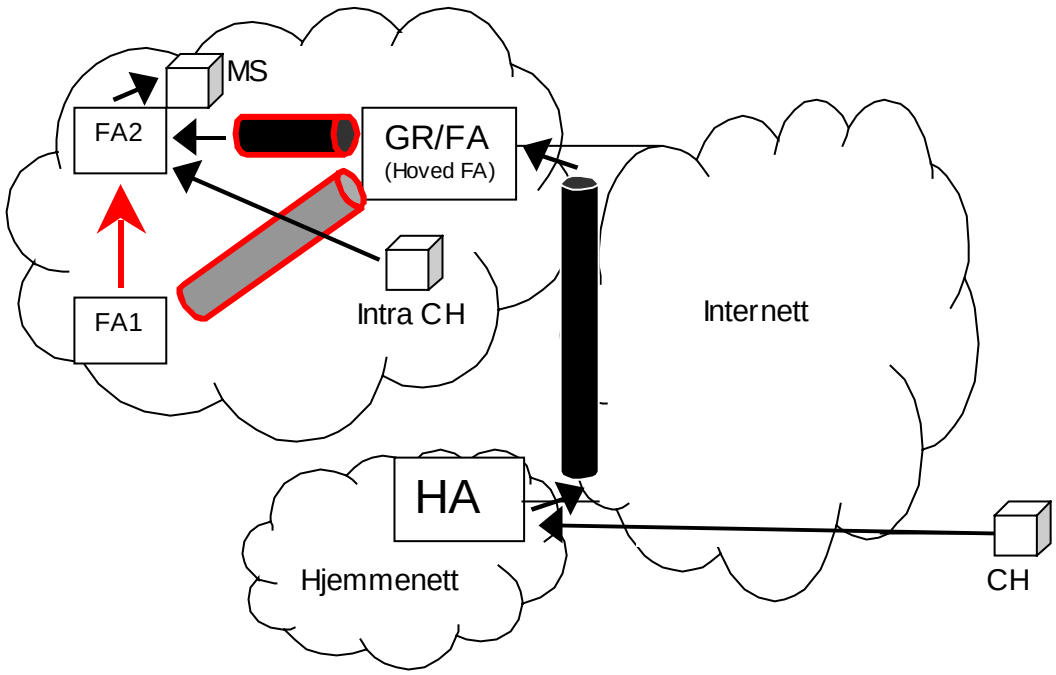
- En global COA som mobilnoden har hos en hoved fjernagent som befinner seg på samme nivå eller sammen med gatewayen i det administrative domenet mobilnoden besøker. Denne adressen er enden for den første tunnelen (markert med 1 på figur 10-2).

Hjemmeagenten vet kun om den globale COA adressen som MS'en har hos hoved FA'en, og derfor blir denne enden på den første tunnelen. Hoved FA'en vet om mobilnodens lokale COA adresse, og vil danne den andre tunnelen som går fra hoved FA'en og som ender hos mobilnodens lokale COA adresse.

Når det gjelder kommunikasjon med korresponderende noden vil disse benytte forskjellige adresser avhengig av deres posisjon. Korresponderende noder som befinner seg utenfor siten vil benytte mobilnodens globale COA, mens noder (Intra CH) som befinner seg innenfor siten vil benytte den lokale COA adressen.
Med denne oppbygningen vil det kunne skilles mellom to forskjellige typer handover: Intra-site handover og intersite handover.

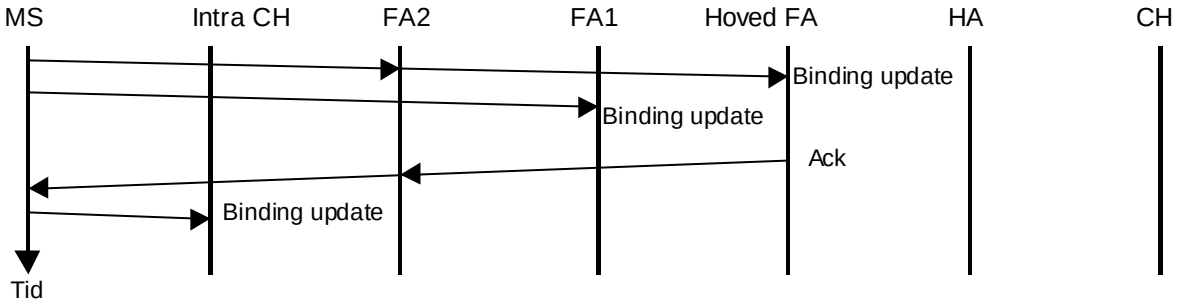
10.2.3 Intra-site handover

Intra-site handover er handover som oppstår når mobilnoden beveger seg innenfor siten, og da er det kun tunnelen fra hoved FA'en og en lokal FA som forandres. Ved slike handover vil det kun være behov for å oppdatering innenfor siten, mens hjemmeagenten og korresponderende noder utenfor siten trenger ingen oppdatering. Dette blir vist i figur 10-3.



Figur 10-3 – Handover innenfor en site

Figur 10-3 viser et nett med et to nivåers hierarki, som betyr at det er to tunneler mellom hjemmeagenten og mobilnoden. Situasjonen her er at mobilnoden beveger seg over i et nytt subnett, fra fjernagent 1 (FA1) og over til subnett 2 hvor FA2 befinner seg. Denne bevegelsen er kun innenfor siten, og det er kun hoved FA'en og korresponderende noder innenfor siten (Intra CH) som må oppdateres (pluss at i tillegg sendes også en binding update til den forrige fjernagenten).
Hjemmeagenten trenger ingen oppdatering da det ikke vil oppstå noen forandringer sett fra dens side, da den globale COA adressen ikke forandres i denne situasjonen. Dette gjelder også for korresponderende noder som befinner seg utenfor siten MS'en besøker, da disse også bruker mobilnodens globale COA adresse.
Registreringsforløpet for denne situasjonen er vist i figur 10-4.

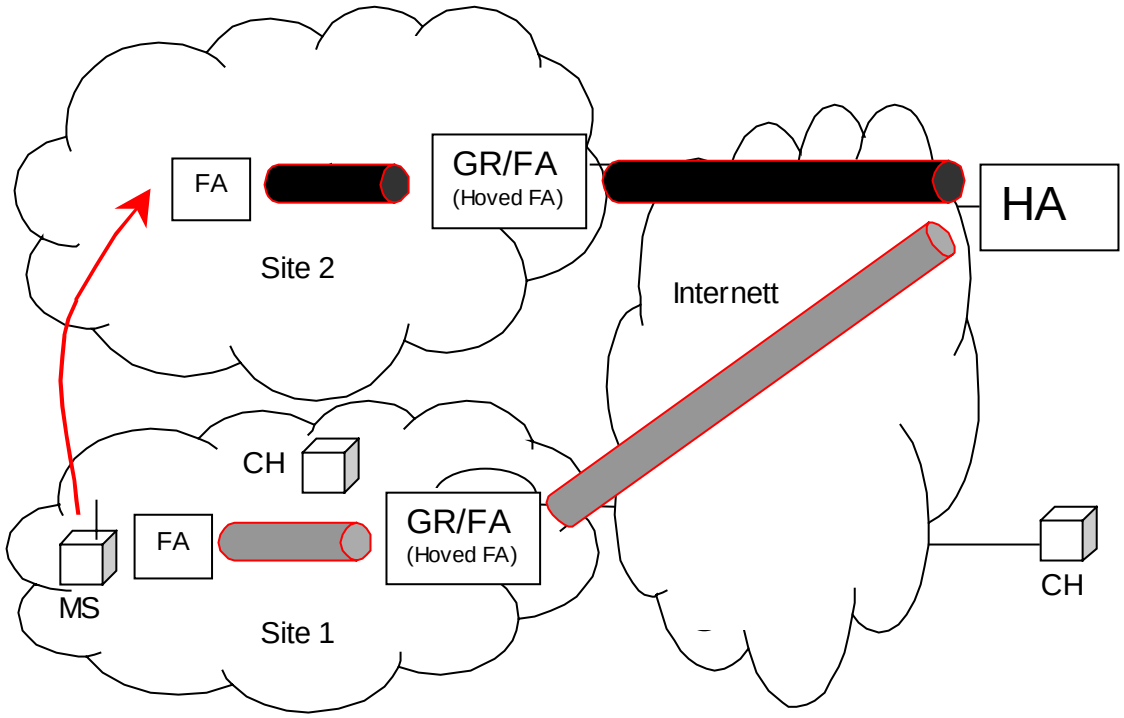


Figur 10-4 – Intra-site Handover

Figur 10-4 viser registreringsforløpet etter at MS'en har fått seg en ny lokal COA adresse hos den nye fjernagenten (FA2). Mobilnoden vil først sende en *Binding Update* gjennom den nye lokale fjernagenten FA2 til hoved FA'en slik at hoved FA'en blir oppdatert med den nye lokale COA adressen. I tillegg vil også mobilnoden sende en BU til den forrige fjernagenten (FA1) slik at trafikk som er på vei til den gamle lokale COA adressen blir videresendt. Til slutt vil mobilnoden sende BU meldinger til evt. korresponderende noder som befinner seg i den samme siten som mobilnoden (intra CH), ettersom disse bruker mobilnoden lokale COA adresse.

10.2.4 Inter-site handover

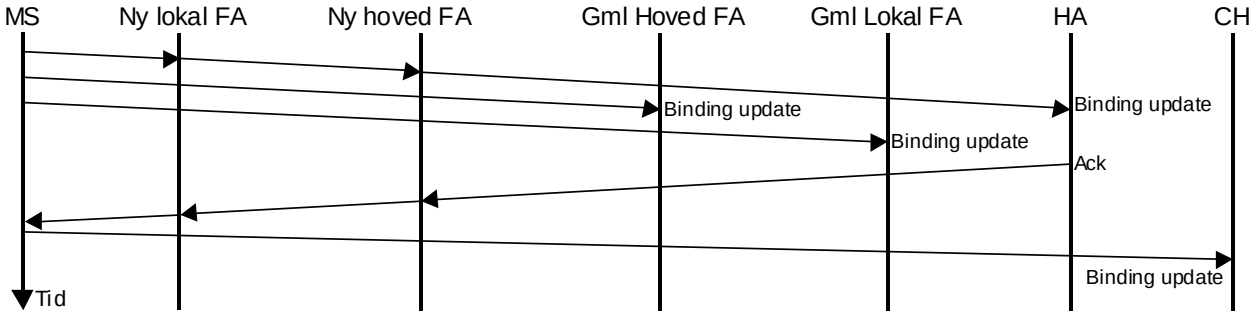
Inter-site handover er som ordet tilsier handover mellom to forskjellige administrative domener, og fører til at begge tunnelene forandres, noe som krever oppdatering hos hjemmeagenten, og alle korresponderende noder. Denne situasjonen blir vist i figur 10-5.



Figur 10-5 – Handover mellom to sites – inter-site handover

Figuren viser en situasjon hvor mobilnoden flytter seg fra site 1 og over til site 2. Ettersom noden her flytter seg mellom to administrative domener, må begge tunnelene forandres, og

mobilnoden må skaffe seg både ny lokal fjernagent, og ny hoved fjernagent. Etter at noden har kommet seg på nett i det nye domenet og skaffet seg nye COA adresser må hjemmeagenten og alle korresponderende noder oppdateres (både de som lå innenfor siten mobilnoden besøkte, og de utenfor). I tillegg oppdateres de gamle fjernagentene for å minske pakketapet.
I situasjoner som denne kan det være vanskelig å få til en såkalt ”soft handover”, ettersom man her er prisgitt forsinkelsen som finnes over Internett for å få registrert de nye COA adressene.
Registreringsforløpet i denne situasjonen er vist i figur 10-6.



Figur 10-6 – Inter-site Handover

Figuren viser registreringsforløpet etter at mobilnoden har kommet seg på nett i det nye administrative domenet (site 2) og skaffet seg nye COA adresser. Det første som skjer er at MS’en sender en binding update til hjemmeagenten via de nye fjernagentene slik at hjemmeagenten kan fjerne den gamle bindingen til site 1, og opprette en ny tunnel til den nye globale COA adressen i site 2. I tillegg til dette sender mobilnoden en BU til de to gamle fjernagentene som den benyttet i site 1 slik at data som er underveis til de gamle COA adressene kan bli fanget opp og videresendt. Når mobilnoden har fått svar og godkjenning fra hjemmeagenten vil mobilnoden sende binding updates til **alle** korresponderende noder.

Når det gjelder tidspunktet for sending av binding updates så viser figur 10-6 at BU meldinger til gamle agenter blir sendt ut før den har fått svar fra hjemmeagenten. Rent ytelsesmessig er dette den beste løsningen, da det er ønskelig at de gamle agentene får oppdatering om lokasjonsforandring så tidlig som mulig slik at pakker kan bli videresendt til den nye globale COA adressen og minst mulig pakker går tapt. Et mulig problem med dette er at BU meldingene blir sendt før den nye COA adressen er godkjent av hjemmeagenten, noe som er et tema som det er diskusjon og tvil om innenfor Mobil IP arbeidsgruppen om dette skal være lovlig.

10.2.5 Oppsummering

Ved å benytte en hierarkisk løsning som skiller mellom forskjellige typer handover vil man kunne oppnå flere forbedringer i forhold til dagens Mobil IP:

- Når mobilnoden beveger seg innenfor et administrativt domene, vil handover ytelsen forbedres (registreringsprosessen vil ta kortere tid), ettersom mobilitetsagentene som må oppdateres befinner seg mye nærmere og forsinkelsen over nettet er dermed kortere.

- En annen fordel er at når mobilnoden beveger seg innenfor en site, er det ikke behov for å oppdatere hjemmeagenten og evt. korresponderende noder utenfor siten, noe som vil minske trafikkbelastningen på Internett forårsaket av Mobil IP.

Det er kun når mobilnoden beveger seg fra en site til en annen at en fortsatt kan få problemer, men dette burde ikke bli noe stort problem, da undersøkelser har vist at 69% av en mobil brukers lokasjonsforandring er lokale forflytninger [12, side 3].
Hvis Mobil IP skal inneha funksjonaliteten som beskrevet her vil det kreve en del forandringer både på klientsiden (mobilnoden) og serversiden (agentene) i Mobil IP. Hos mobilnoden må det blant annet innføres mekanismer for få håndtere flere COA adresser, og hos fjernagentene må agent kunngjøringene forandres, slik at mobilnoden er i stand til å få tak i flere COA adresser.

10.2.5.1 Skalering

Med Mobil IP slik den er spesifisert i dag vil mengden av handoverregistreringer i framtiden kunne føre til betydelig trafikkbelastning på Internett når antallet mobile noder blir stort. Grunnen til dette er mobilitet håndteres likt uansett om det er små bevegelser innenfor en site, eller globale bevegelser. Dette vil kunne gi dårligere ytelse når en betydelig andel av trafikken i nettet er handoverregistreringer. Flaskehalsene her vil ligge i nettene, og ikke hos agentene da deres prosesseringskapasitet sjelden vil være et problem [29].
Løsningen med hierarkisk oppbygning vil også kunne bøte på dette, da lokale forflytninger ikke vil føre til behov for oppdatering hos hjemmeagenten og eksterne korresponderende noder. Belastningen på nettene vil derfor bli mindre, og gjøre Mobil IP bedre skalerbar i større sammenhenger.
Denne løsningen har også gode skaleringsmuligheter da ytelsen kan forbedres ytterligere ved å øke antall nivåer i hierarkiet, slik at lasten forårsaket av Mobil IP kan deles på flere, og avstandene mellom agentene minskes. (Antall nivåer blir bestemt ut i fra størrelsen på nettet, trafikkmengde og antallet mobile brukere). Ved å minske avstanden mellom agentene vil forsinkelsen synke, og handover kunne foregå raskere.

10.2.6 IPv6

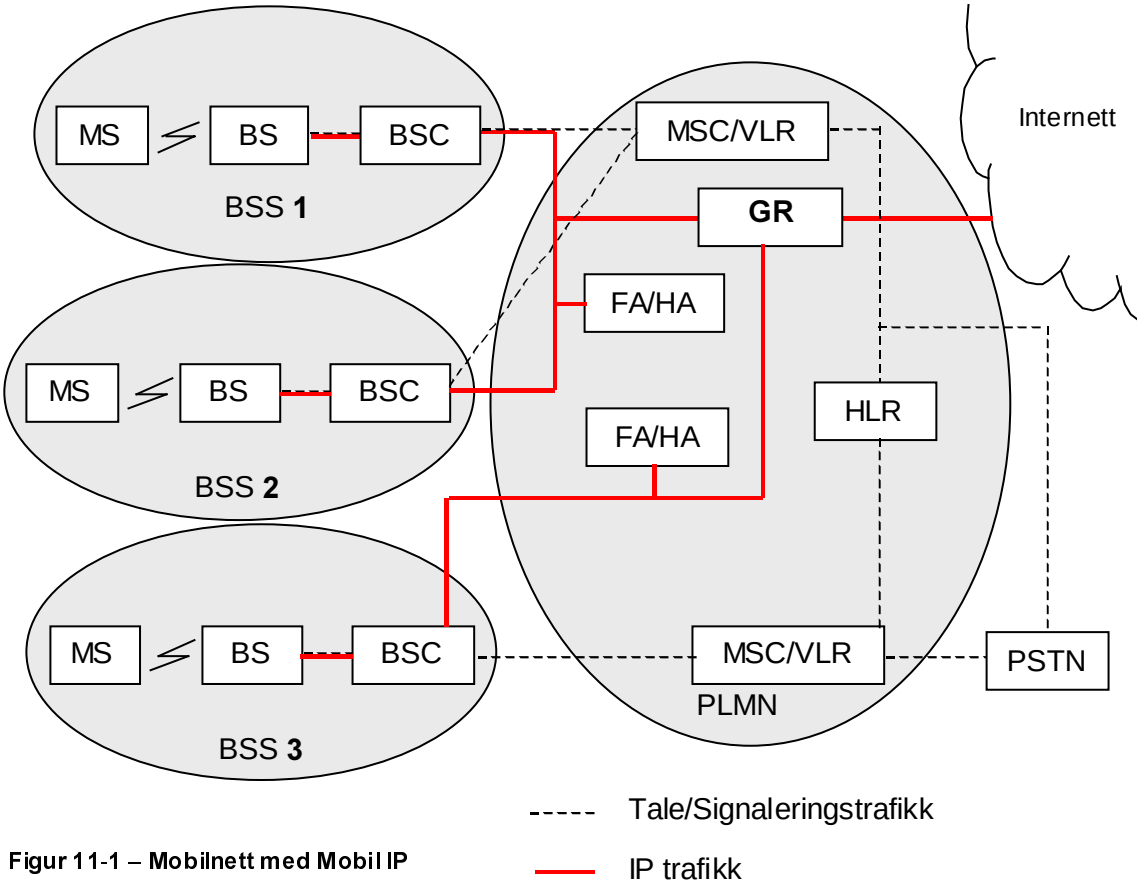
I IPv6 fungerer Mobil IPv6 stort sett på samme måte som i IPv4, og de samme forbedringene vil også kunne gjelde her. Det er en vesentlig forskjell i IPv6 som er at i Mobil IPv6 finnes det ingen fjernagent. I Mobil IPv6 er det meningen at mobilnoden skal klare seg uten hjelp i det hele tatt, og bare operere slik som med co-located COA i Mobil IPv4. For å kunne bygge opp et system med hierarki må det innføres en type mobilitetsagenter også i besøksnettene, slik at det er mulig med flere tunneler mellom mobilnoden og hjemmeagenten. Uten dette vil det ikke være mulig forbedre handover på denne måten, og minske nettbelastningen på Internett forårsaket av Mobil IP [12]. En løsning som ikke skiller mellom forskjellige typer bevegelser slik som Mobil IPv6 er spesifisert i dag vil være lite skalerbar og kunne fungere dårlig i stort omfang.
Løsningen blir derfor å innføre en form for fjernagenter også i Mobil IPv6 for å muliggjøre de samme fordelene som er foreslått for Mobil IPv4.

11 Mobil IP i GPRS

Ved å bruke Mobil IP som mobilhåndteringssystem i framtidens mobilsystemer kan en oppnå flere fordeler: Brukeren vil være i stand til å bevege seg mellom forskjellige nett basert på forskjellige teknologier uten problemer, mobilnoden kan roame, dvs. besøke andre nett som er basert på andre teknologier enn GSM/GPRS. Med Mobil IP er en ikke avhengig av kun en type teknologi slik som mobilnettene er i dag, men mobilnettleverandøren kan også benytte andre teknologier som trådløse LAN, for å gi brukerne høyere ytelse og bedre tjenester i enkelte områder. Med kun bruk av GPRS vil en proprietær teknologi brukes, og som bare kan brukes sammen med GSM, og tilbyr på ingen måte slik fleksibilitet som Mobil IP vil kunne gjøre.

11.1 Forslag til oppbygning

Hvis Mobil IP skal brukes som mobilhåndteringssystem i GPRS, kan nettet bygges opp ved at GSN nodene i GPRS fjernes, byttes ut med tradisjonelle IP rutere og Mobil IP mobilitetsagenter. GPRS spesifikke løsninger blir fortsatt brukt i basestasjonsubsystemet, dvs. fra BSC'en og utover. BSC'en blir derfor grensesnittet til det interne IP nettet i mobilnettet, og må derfor inneha IP funksjonalitet (det har den ikke slik GPRS er spesifisert i dag). For GPRS delen er det mobilitetsagentene (fjernagenter og hjemmeagenter) som tar seg av de funksjonene som GSN og HLR nodene har i dag.
Ved å benytte Mobil IP sammen med GPRS, kan nettet bygges opp som vist i figur 11-1 [19].



Figur 11-1 – Mobilnett med Mobil IP

Figur 11-1 viser et nett, hvor GSN nodene i GPRS er byttet ut med et IP nettverk med rutere og mobilitetsagenter. Vanlig tale, linjesvitsjet data og signaleringstrafikk går fra BSC til de samme nodene som før (MSC/VLR og HLR). BSC'ene vil med denne arkitekturen også inneha IP funksjonalitet og funksjoner mellom MS'en og IP nettet. Nettet er ellers bygd opp slik som et standard IP nettverk med mobilitetsagenter i hvert subnett. Alle subnett innenfor mobilnettet inneholder både fjernagent og hjemmeagenter, og det er agenten i det første subnettet mobilnoden logger seg på som blir dens hjemmeagent. Når mobilnoden beveger seg rundt i mobilnettet, vil alle agenter i andre subnett være fjernagenter.

11.1.1 Handover

Når mobilnoden beveger seg rundt vil det kunne forekomme 3 forskjellige typer handover med en slik oppbygning:

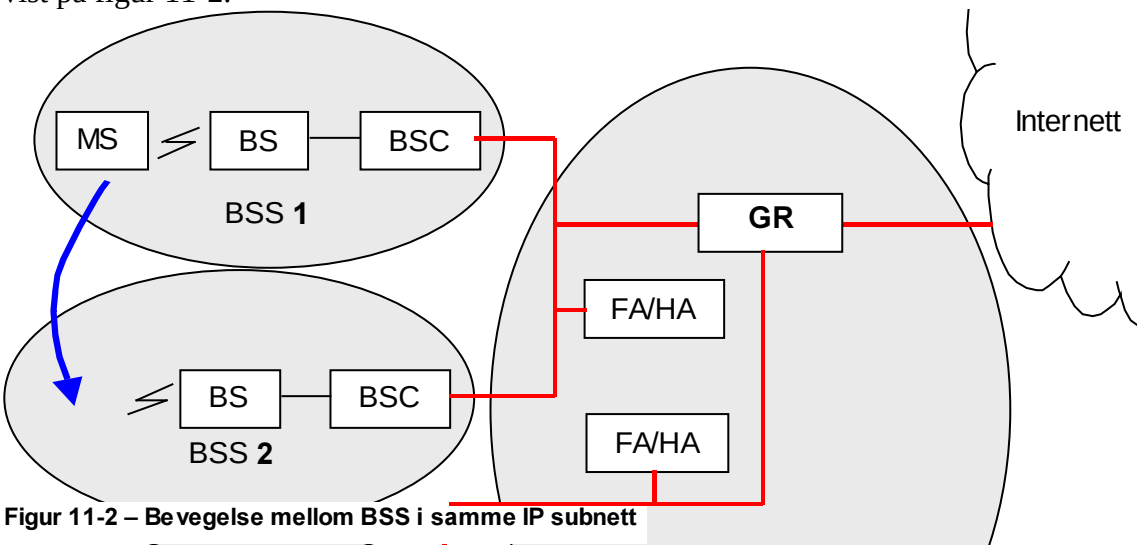
- Bevegelse mellom basestasjoner innenfor én BSC.
- Bevegelse mellom basestasjoner som tilhører forskjellige BSC'er, men begge BSC'er er i samme IP subnett.
- Bevegelse mellom basestasjoner som tilhører forskjellige BSC'er, men hvor BSC'ene befinner seg i forskjellig IP subnett.

11.1.1.1 Handover innenfor samme BSC

Hvis mobilnoden beveger seg mellom basestasjoner innenfor en BSC vil den kunne bevege seg innenfor samme subnett, og det kreves ingen oppdatering noe sted på IP nivå. Situasjonen tilsvarer en horisontal handover, hvor det kun er lavere lags (GSM/GPRS) protokoller er involvert og håndterer handoveren.

11.1.1.2 Handover mellom to BSC'er i samme subnett

Det andre tilfellet hvor mobilnoden beveger seg til en basestasjon som tilhører en annen BSC, men som befinner seg innenfor samme IP subnett for eksempel mellom BSS 1 og BSS 2 som vist på figur 11-2.



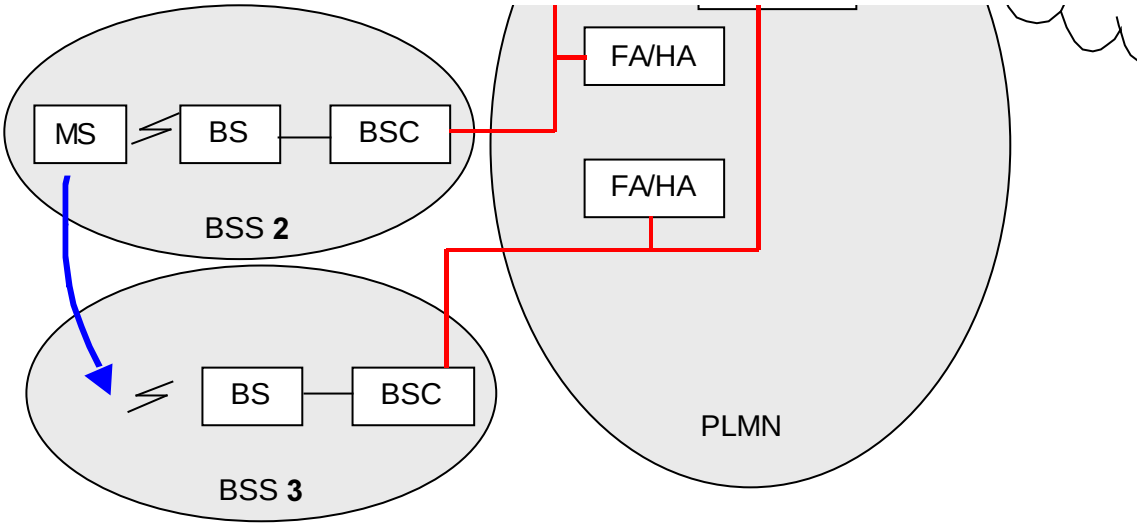
Figur 11-2 – Bevegelse mellom BSS i samme IP subnett

Her vil det ikke være behov for noen oppdatering på Mobil IP nivå, og selve handoveren vil foretas av lavere (GSM/GPRS) protokoller. Det eneste som må til er kun en oppdatering slik at ruteren sender dataene til riktig BSC. Dette gjøres ved å benytte ARP [35], som har en proxy mekanisme som gjør at en node kan ta i mot data som egentlig ikke er adressert til den selv. På denne måten kan BSC'en ta i mot pakker på vegne av mobilnoden. Alle mobilnoder

som befinner seg under en BSC vil ha hver sin COA IP adresse, men i ARP cachen i ruteren vil de være linket til samme lavere lags adresse. Det som må gjøres når mobilnoden beveger seg til en annen BSC i samme subnett, er at den nye BSC'en vil sende ut en *unsolicited ARP* reply (tilsvarende *proxy neighbour advertisement* i IPv6) for å fortelle ruteren at den aktuelle IP adressen (dvs. mobilnodens lokale COA) har byttet til en annen laverelags adresse.

11.1.1.3 Handover mellom to BSC'er i forskjellige subnett

Hvis mobilnoden beveger seg over til en ny basestasjon som tilhører en BSC i et annet subnett, vil det være behov for en oppdatering i Mobil IP. Grunnen til dette er at mobilnoden må få tak i en ny COA adresse, og få oppdatert denne hos hjemmeagenten. Denne situasjonen er vist i figur 11-3.



Figur 11-3 – Bevegelse mellom BSS i forskjellige IP subnett

Handoveren her vil bli helt tilsvarende en vanlig Mobil IP handover som fortalt i kap. 6, og hvor hjemmeagenten og evt. korresponerende noder trenger oppdatering om den nye COA adressen.

12 Forslag til arkitektur med Mobil IPv6

Et eksempel på et scenario med anvendelse av Mobil IPv6 i fremtiden kan være at en mobilnettleverandør kan bli en slags trådløs tilbyder, som i tillegg til vanlig tale og datatjenester kan tilby høyhastighetstilgang til Internett for sine kunder. Med en løsning med Mobil IP kan tilbyderen bygge opp et nett benytte flere teknologier for å gi best mulig ytelse. Nettet kan for eksempel være basert på bruk av HIPERLAN, 802.11 Trådløst LAN, og GSM med GPRS. Steder hvor det er behov og et sannsynlig marked for høyhastighets nett, som for eksempel flyplasser, togstasjoner, og bussterminaler kan høyytelsesnett som HIPERLAN installeres, mens generelt i godt befolkede områder kan være dekket av 802.11 trådløse LAN for å gi god ytelse. For utkantområder hvor markedet og behovet er mindre vil GSM med GPRS være tilgjengelig. I et slikt scenario hvor Mobil IP skal kunne brukes i fremtidens mobilsystemer må den tilfredsstille en rekke krav når det gjelder handoverytelse. I tillegg må det også finnes systemer som tar seg av AAA tjenester (Authentication, Authorization, Accounting).
Dagens Mobil IP (v4 og v6) er ikke i stand til å tilfredsstille disse kravene, men hvis en får forbedret handover ytelsen og skaleringsmulighetene slik som foreslått i kapittel 10 vil Mobil IP standarden kunne gi god ytelse. For å tilrettelegge dette med AAA funksjonalitet kan det være fordelaktig å benytte såkalt NAI (*Network Access Identifier*) [31] i stedet for den faste IP adressen som identifikator i Mobil IP. Ellers når det gjelder AAA funksjonalitet så blir det videre gått ut fra at dette eksisterer og fungerer, men på grunn av avgrensningen av oppgaven er ikke dette sett nærmere på.

12.1 Network Access Identifier

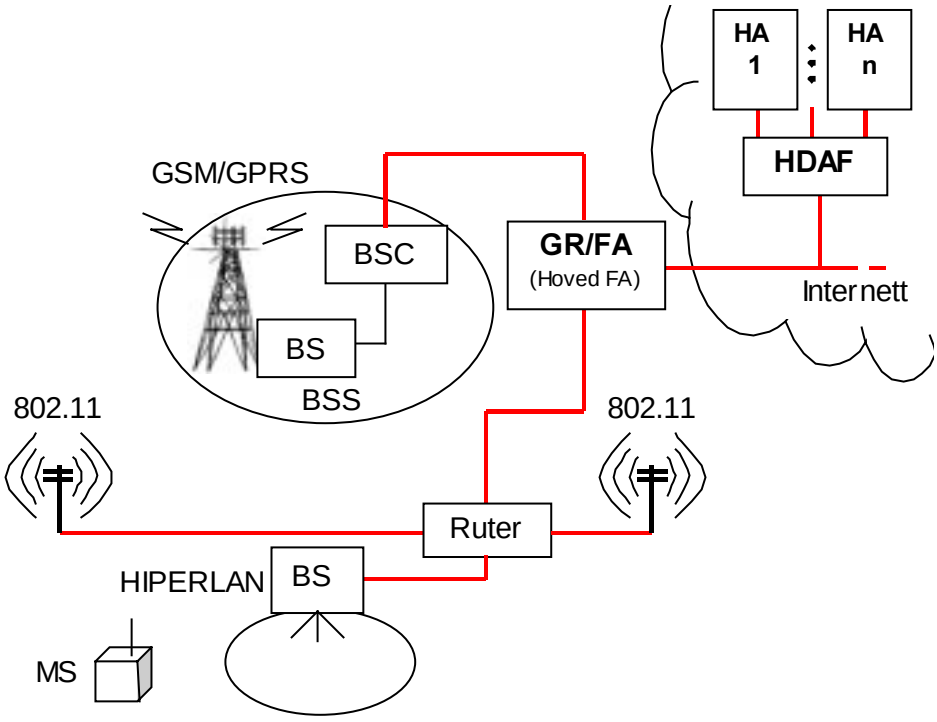
Bruk av NAI, som tillegg eller erstatning av den faste IP adressen som identfikator er blitt foreslått til Mobil IP hvor NAI benyttes i stedet [30]. Dette forslaget går ut på å legge til et NAI tilleggsfelt i extension feltet i Mobil IP reg. request meldingene. MN-NAI feltet som det kalles vil inneholder en unik identifikator, som fungerer som en userID og benyttes til autentisering og identifisering. Identifikatoren i NAI feltet er formatert omtrent som en vanlig e-mail adresse: navn@domene, hvor navn er brukeren/mobilnoden, og domenet er ISP'en (*Internet Service Provider*) som noden er abonnent hos.
Opprinnelig var NAI tiltenkt til bruk for identifisering og autentisering av brukere i tunnelleringssammenheng (VPN – *Virtual Private Networks*), og i roaming slik at brukere skal kunne benytte seg av flere ISP'er. Senere ble også NAI tatt i bruk i AAA sammenheng for AAA servere i Internett (som RADIUS og DIAMETER) for identifikasjon av brukere.

12.1.1 NAI i Mobil IP

Ved bruk av NAI i Mobil IP sammenheng blir det enklere å innføre AAA funksjonalitet, og mobilnoden kan bli identifisert og autorisert uten å ha noen fast hjemmeadresse. I stedet for en fast hjemmeadresse kan mobilnoden som identifikasjon for eksempel benytte IMSI-nummeret (*International Mobile Station Identifier*) som mobilnoden allerede har fra før i GSM/GPRS sammenheng. For situasjoner hvor mobilnoden ikke lenger har noen fast hjemmeadresse finnes det i [30] et forslag som går ut på å legge til en ekstra node, en såkalt *Home Domain Allocation Function* (HDAF) node som kan tildele hjemmeadresser og hjemmeagent dynamisk til mobilnoder.

12.2 Forslag til arkitektur

Ved å benytte Mobil IPv6 sammen med flere trådløse teknologier, bruk av HDAF node og NAI kan et nett bygges opp som vist på figur 12-1. Scenarioet er også basert på forbedringsforslagene fra kap. 10.



Figur 12-1 – Scenario med bruk av Mobil IPv6

Figuren viser et IPv6 nettverk hvor flere teknologier blir brukt: GSM med GPRS, 802.11 WLAN, og HIPERLAN. Nettverkstopologien er bygd opp slik at den ene teknologien har dekningsområde som dekker den andre som beskrevet i kap. 9.2.1. Hos mobilnoden er funksjonaliteten som beskrevet i kap 10.1 implementert når det gjelder vertikale handover. Scenarioet tar også i bruk den hierarkiske oppbygningen som er foreslått i kap. 10.2, hvor gateway ruteren har rolle som hoved fjernagent.
Når en mobilnode logger seg på for første gangen i mobilnettet, vil den benytte NAI identifikatoren for å få tildelt en hjemmeagent og hjemmeadresse hos HDAA'en, og skaffe seg to COA adresser. Mobilnoden vil få tak i en lokal COA adresse i subnettet den er i enten via *stateless autoconfiguration* eller med DHCPv6, og en global COA adresse hos GR/FA noden. Den globale COA adressen vil bli brukt for tunnelen fra hjemmeagent til hoved FA, og som mobilnodens COA adresse for korresponderende noder som befinner seg utenfor mobilnettet dvs. i Internett. Den lokale COA adressen vil bli brukt for tunnel fra hoved FA'en til mobilnoden, og for evt. korresponderende noder som befinner seg innenfor mobilnettet.

12.2.1 Utredning av scenario

Et spørsmål som kan stilles i dette scenarioet er kanskje hva er formålet med en slik løsning? Om få år (2002) vil UMTS bli lansert, og den vil kunne tilby mye av den samme funksjonaliteten. Derfor kan denne løsningen til en viss grad ses på som en konkurrent til UMTS, da den vil tilby mange av de samme tjenestene:

- tradisjonelle taletjenester gjennom GSM delen.
- høyhastighetsdataforbindelser over GPRS og trådløse LAN.

Et motargument for å benytte trådløse LAN teknologier (802.11 og HIPERLAN) i stor skala er at de er da de er basert på aksessmetoder hvor mediet er delt mellom alle brukerne, og det finnes ingen allokering av ressurser. Derfor vil det i situasjoner med mange brukere kunne gi dårlig ytelse på grunn av for mange brukere og mye kollisjoner.
I tillegg har disse teknologiene problemer med skjulte noder (på grunn av deres *carrier sense* aksessmetode), noe som ikke er et problem i et dedikerte mobilnett da disse er bygd opp på en måte hvor dette har blitt tatt hensyn til fra begynnelsen av. De trådløse LAN teknologiene som er benyttet i eksempelet her har heller ingen muligheter for garantert tjenestekvalitet, de har kun en slags prioritetsmekanisme for pakker, men kan ikke garantere noen ting, og de tilbyr kun en ”best effort” tjeneste.
Når det gjelder utnyttelse av frekvenser er ikke dette helt optimalt på trådløse LAN, da alle noder og basestasjoner bruker samme frekvenser, og med frekvenshopping (802.11) for at naboceller skal forstyrre hverandre minst mulig. For å forbedre kapasiteten uten å forandre standardene (frekvensbruk, modulering osv.) kan sektorisering av dekningsområdene benyttes ved bruk av direktive antenner i stedet for de vanlige rundtstrålende antennene som benyttes i dag. Hver sektor blir da et separat trådløst LAN, og for å få best kapasitet bør hver antenne/aksess punkt være tilknyttet et backbone nett med svitsjet arkitektur.

De problemene som er fortalt her for trådløse LAN vil sannsynligvis ikke være et problem med UMTS, fordi UMTS er et nytt system som blir bygd opp fra grunnen av, og slike ting blir derfor tatt hensyn til. UMTS har også sine begrensninger, blant annet en maksimum hastighet på kun 2Mbit/s i første fase (som kommer i 2002).
En stor fordel med Mobil IP arkitekturen er at hvis en ønsker det kan en slik løsning faktisk realiseres allerede i dag i mindre skala, hvis en aksepterer problemene med vertikale handover, og unngår store nett på grunn av den dårlige skaleringen. Allerede i løpet av år 2000 vil Mobil IP kunne inneha funksjonalitet tilsvarende det som er beskrevet i denne rapporten, noe som er mye tidligere enn lanseringen av UMTS.

13 Diskusjon & Konklusjon

I denne hovedoppgaven har jeg presentert problemstillingen rundt handover i Mobil IP, og kommet med forslag til forbedringer. I tillegg er det i rapporten også presentert teknologiene som er viktige for forståelsen av problemstillingen og for anvendelsen av Mobil IP.

Håndteringen av handover i Mobil IP i dag er ikke god nok for en fremtidig plattform for mobilitet. Avbruddet og pakketapet forårsaket av handover er så stort at ytelsen i noen situasjoner vil bli betraktelig redusert for høyerelags protokoller og realtime applikasjoner blir ubrukelige. Mobil IP har i dag heller ingen intelligent mekanisme for håndtering av flere nettverksgrensesnitt (vertikale handover), og kan ikke se forskjell på forskjellige typer bevegelser. Om mobilnoden beveger seg lokalt, regionalt, eller globalt har ingen betydning, og hjemmeagenten og alle korresponderende noder må oppdateres uansett. Dette vil gi Mobil IP dårlige skaleringsegenskaper, og hvis Mobil IP blir implementert i stor skala slik den er i dag vil det oppstå betydelig trafikkbelastning på nettet på grunn av handovermeldinger.

Handoverproblemstillingen i Mobil IP er et tema som arbeides og forskes på en rekke steder i dag, og det finnes derfor mange forslag til forbedringer tilgjengelig. Mine forslag til forbedringer bygger til en viss grad på en del av disse resultatene, ved at jeg har benyttet deler av dem, satt de sammen på forskjellige måter, og i tillegg kommet med egne forslag til forbedringer.
Jeg har i rapporten sett på handoversituasjonen oppdelt i de forskjellige fasene mobilnoden går i gjennom når handover foretas, og har sett på hver fase hver for seg, og hva som kan forbedres der.

Fasene som mobilnoden går i gjennom er hovedsakelig:
- Oppdagelse av nytt nett.
- Oppnå ny COA adresse.
- Lokasjonsoppdatering, dvs. registrering av den nye COA adressen.

Hovedmålet er handoverforløpet skal ta så kort tid som mulig, og med minst mulig pakketap. Det første punktet går på vertikale handover, som kan forbedres ved utnyttelse av linklags informasjon, og benytte *agent advertisement* meldingene slik at mobilnoden skjønner raskere at den har mistet forbindelsen, og bruke kortere tid på å få tak i en ny agent.

Når det gjelder COA adressen er det ønskelig at denne fås tak i hurtigst mulig, og den hurtigste måten er å bruke en fjernagents IP adresse som COA, noe som også har den fordelen at flere noder deler den samme COA adressen, og derfor fører til mindre belastning på det begrensede adresserommet i IPv4.

Den siste fasen er oppdatering av den nye COA adressen, og kan forbedres ved å bygge opp nettene med et hierarki av mobilitetsagenter, slik at det kan skilles mellom forskjellige typer bevegelser. Med en hierarkisk oppbygning vil handover kunne ta kortere tid ettersom avstanden til agenten(e) som må oppdateres kan bli mye kortere. Denne oppbygningen vil også kunne gjøre Mobil IP bedre skalerbar, når ikke alle noder må oppdateres for hver eneste lokasjonsforandring.

Til sammen vil disse forbedringene kunne gjøre at handover vil kunne foretas mye raskere enn i dag, og i enkelte situasjoner helt uten pakketap.
Forslagene til forbedringer presentert i denne rapporten vil etter min mening gjøre handoversituasjonen i Mobil IP mye bedre, men kan sannsynligvis forbedres ytterligere,

muligens ved anvendelse av anycast og multicast i en eller annen sammenheng noe som det finnes enkelte andre forslag på.

Hvis Mobil IP skal kunne bli en slags felles plattform for mobile tjenester i fremtiden er det mye arbeid som gjenstår, men hvis forbedringene som er foreslått i denne rapporten blir implementert i Mobil IP vil handoversituasjonen bli mye bedre, og vil nærme seg de krav som vil gjelde for fremtiden.
Da det ikke er kun handover som er problemet med Mobil IP, må også en del andre områder forbedres før en kommer i mål:

- Implementering av AAA tjenester (NAI)
- Forbedre skalering (fordeling av last hos agenter, minimalisere trafikk forårsaket av Mobil IP).
- Gi muligheter for tjenestekvalitet hos mobile noder.
- Forbedre samspillet med mobile nett (for eksempel GPRS og UMTS).

Jeg mener at Mobil IP har et godt potensiale til å kunne bli fremtidens standard for mobile tjenester, men det gjenstår mye arbeid før vi er der. Mobil IP plattformen kan da bli en farlig konkurrent til UMTS, ettersom den har en rekke fordeler: den er plattformuavhengig, er uavhengig av underliggende teknologier (og kan derfor benytte seg av nye teknologier mye raskere), den er transparent for høyere lags protokoller og applikasjoner, og er en åpen standard som kan brukes av alle.
Resultatet kan bli at det er Mobil IP som blir framtidens standard for mobilitet, og UMTS blir mer sett på som en aksessteknologi på lik linje med GSM/GPRS og Wireless LAN.

14 Litteraturreferanser

[1] Larry L. Peterson, Bruce S. Davie:
***Computer Networks – A Systems Approach**, Morgan Kaufmann Publishers, 1996.*

[2] William Stallings:
***IPv6: The New Internet Protocol**, IEEE Communications Magazine page 96-108, July 1996.*

[3] Stewart S. Miller:
***IPv6 The Next Generation Internet Protocol**, Digital Press, 1998,*
http://www.bh.com

[4] Mark A. Miller:
***Implementing IPv6 – Migrating to the Next Generation Internet Protocols**, M&T Books, 1998.*

[5] IETF Network Working Group:
***Reserved IPv6 Subnet Anycast Addresses**, RFC (Request For Comments) – 2526, IETF, March 1999,*
http://www.ietf.org/html.charters/ipngwg-charter.html.

[6] IETF Network Working Group:
***Neighbour discovery for IP version 6**, RFC 2461, IETF, December 1998,*
http://www.ietf.org/html.charters/ipngwg-charter.html.

[7] Charles E. Perkins:
***Mobile IP - Design Principles and Practises**, Addison Wesley Wireless Communication Series (1998),*
http://www.awl.com.

[8] Charles E. Perkins, Sun Microsystems:
***Mobile IP**, IEEE Communications Magazine page 84-99, May 1997.*

[9] IETF Mobile IP Working Group:
***Mobility Support in IPv6**, internet draft, IETF, November 1998,*
http://www.ietf.org/html.charters/mobileip-charter.html.

[10] William Woo, Victor C.M. Leung:
***Handoff Enhancement in Mobile IP enviroment**, Dept. of Electrical Engineering, University of British Columbia & Motorola Wireless Data Grop page 760-764, 1996.*

[11] IETF Mobile IP Working Group:
***Special Tunnels for Mobile IP**, internet draft, IETF, November 1997,*
http://www.ietf.org/html.charters/mobileip-charter.html.

[12]Claude Castelluccia:
***A Hierarchical Mobile IPv6 Proposal**, INRIA, November 1998.*

[13]Charles E. Perkins:
***Mobile Networking in the Internet**, Mobile networks in the Internet, Mobile networks & applications, Baltzer science publisher & ACM.*

[14]Mark Stemm & Randy H. Katz:
***Vertical Handoffs in wireless overlay networks**, Mobile networks in the Internet, Mobile networks & applications, Baltzer science publisher & ACM.*

[15]Ramon Caceres & Venkata N. Padmanabhan:
***Fast and scalable wireless handoffs in support of mobile Internet audio,** Mobile networks in the Internet, Mobile networks & applications, Baltzer science publisher & ACM.*

[16] Jost Weinmiller, Morten Schläger, Andreas Festag, Adam Wolisz:
***Performance study of access control in wireless LANs – 802.11 DFWMAC and ETSI RES 10 Hiperlan**, Mobile Networks and Applications 2 page 55-67, Baltzer Science Publisher BV, 1997.*

[17] Richard O. LaMaire, Arvind Krishna, and Pravin Bhagwat, IBM, James Panian, Ericsson Inc:
***Wireless LANs and Mobile Networking: Standards and Future Directions**, IEEE Communications Magazine page 86-94, August 1996.*

[18] Brian P. Crow, Indra Widjaja, Jeong Geun Kim, Prescott T. Sakai:
***IEEE 802.11 Wireless Local Area Networks**, IEEE Communications Magazine page 116-126, September 1997.*

[19] Russel Hsing:
*Final Report on **PCS-to-Internet Protocol Interworking**, Broadband Wireless Networking and Technology, Bellcore, June 24, 1998.*

[20] Mobile IP working group:
***Requirements on Mobile IP from a Cellular Perspective, internet draft**, IETF, February 1999, http://www.ietf.org/html.charters/mobileip-charter.html.*

[21] Jian Cai, David J. Goodman, Rutgers University:
***General Packet Radio Service in GSM**, IEEE Communications Magazine page 122-131, October 1997.*

[22] Götz Brasche, Bernhard Walke Aachen University of Technology:
***Concepts, Services and Protocols of the New GSM Phase 2+ General Packet Radio Service**, IEEE Communications Magazine page 94-104, August 1997.*

[23] Gabriela Grolms:
***The role of Internet technology in UMTS**, Project I, Telenor R & D, 1999, http://pi.nta.no*

[24]Gustafsson, Herlitz, Jonsson, Korling:
***UMTS/IMT-2000 and Mobile IP/DIAMETER harmonization**, IETF, November 1998, http://search.ietf.org/internet-drafts/draft-gustafsson-mobileip-imt-2000-00.txt*

[25] IETF Mobile IP Working Group:
***IP Mobility Support**,RFC 2002 , IETF, November 1998, http://www.ietf.org/html.charters/mobileip-charter.html.*

[26] IETF IPNG (ipngwg) Working Group:
***Internet Protocol Version 6 (IPv6) Specification**, RFC 2460, December 1998, http://www.ietf.org/html.charters/ipngwg-charter.html*

[27] IETF Mobile IP Working Group:
***Route Optimization in Mobile IP**, Internet draft, February 1999 http://www.ietf.org/html.charters/mobileip-charter.html.*

[28] IETF Mobile IP Working Group:
***Reverse Tunneling for Mobile IP**, RFC 2344, May 1998 http://www.ietf.org/html.charters/mobileip-charter.html.*

[29] Frederic Jacot:
***Scaling of Mobile IP**, Telenor FoU notat, Mai 1999.*

[30] IETF Mobile IP Working Group:
***Mobile IP Network Address Identifier Extension**, Internet draft, February 1999 http://www.ietf.org/html.charters/mobileip-charter.html.*

[31] IETF Network Working Group:
***The Network Access Identifier**, RFC 2486, January 1999 http://www.ietf.org/.*

[32] IETF Network Working Group:
***Security Architecture for the Internet Protocol (IPsec)**, RFC 2401, November 1998 http://www.ietf.org/html.charters/ipsec-charter.html.*

[33] IETF DHC Working Group:
***Dynamic Host Configuration Protocol for IPv6 (DHCPv6)**, Internet draft, February 1999*
http://www.ietf.org/html.charters/dhc-charter.html.

[34] IETF DHC Working Group:
***Dynamic Host Configuration Protocol**, RFC 2141, March 1997*
http://www.ietf.org/html.charters/dhc-charter.html.

[35] Network Working Group, David C. Plummer:
***Address Resolution Protocol**, RFC 826, November 1982.*

15 Stikkordliste

AAA	Authentication, Accounting & Authorization.
ARP	Address Resolution Protocol
CH	Correspondent Host
COA	Care Of Address
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
DHCP	Dynamic Host Configuration Protocol
ETSI	European Telecommunications Standards Institute
FTP	File Transfer Protocol
FA	Foreign Agent
GPRS	General Packet Radio Service
GSM	Global System for Mobile communications
HA	Home Agent
HTTP	HyperText Transfer Protocol
ICMP	Internet Control Message Protocol
IETF	Internet Engineering Task Force
IEEE	Institute of Electrical and Electronics Engineers
IMSI	International Mobile Station Identifier
IR	Infra Red
ISO	International Organization for Standardization
ISP	Internet Service Provider
LAN	Local Area Network
LLC	Logical Link Control
MAC	Medium Access Control
MTU	Message Transfer Unit
OSI	Open Systems Interconnection

QoS	Quality of Service
RARP	Reverse Address Resolution Protocol
RF	Radio Frequency
SMTP	Simple Mail Transfer Protocol
TCP	Transmission Control Protocol
TDMA	Time Division Multiple Access
UDP	User Datagram Protocol
UMTS	Universal Mobile Telecommunications System
WAN	Wide Area Network
WLAN	Wireless Local Area Network
WWW	World Wide Web