Suoraviestintää ja läsnäoloa SIP:illä

Petri Lintula

Tampereen yliopisto
Tietojenkäsittelytieteiden laitos
Tietojenkäsittelyoppi
Pro gradu -tutkielma
Huhtikuu 2004

Tampereen yliopisto
Tietojenkäsittelytieteiden laitos
Tietojenkäsittelyoppi
LINTULA, PETRI: Suoraviestintää ja läsnäoloa SIP:illä
Pro gradu -tutkielma, 53 sivua
Huhtikuu 2004

SIP eli Session Initiation Protocol on signalointiprotokolla, joka mahdollistaa istuntojen luomisen, muokkaamisen ja lopettamisen ip-pohjaisissa verkoissa. Sitä voidaan käyttää muun muassa suoraviestintään eli käyttäjien väliseen lähes reaaliaikaiseen viestien välittämiseen. Tällöin on tärkeää tietää käyttäjien läsnäolosta eli halukkuudesta ja mahdollisuudesta kommunikoida.

Suoraviestinnällä ja läsnäololla ei ole olemassa yhtenäistä standardia, vaan kaikki ohjelmistot käyttävät omia ratkaisujaan. Tämän vuoksi niiden ei ole mahdollista toimia keskenään. SIP yrittää olla standardi, jonka avulla eri ohjelmistot voisivat toimia keskenään. Näin käyttäjät voisivat vapaasti valita haluamansa ohjelmiston.

Tutkielmallani pyrin selvittämään, miten SIP toimii suoraviestinnän ja läsnäolon kanssa, soveltuuko se tähän tarkoitukseen sekä mitä mahdollisia ongelmia sen käyttöön saattaa liittyä.

Avainsanat ja -sanonnat: SIP, suoraviestintä, läsnäolo
CR-luokat: C.2.2, C.2.6

Sisällys

1. Johdanto

SIP eli Session Initiation Protocol on signalointiprotokolla, joka mahdollistaa istuntojen (sessions) luomisen, muokkaamisen ja lopettamisen ip-pohjaisissa verkoissa. SIP ei ota kantaa istunnossa välitettävään dataan; sitä käytetään pelkästään istuntojen käsittelyyn. Istunto voi esimerkiksi olla kahden käyttäjän välinen Internet-puhelu tai usean käyttäjän välinen multimediakonferenssi.

Suoraviestintä eli Instant Messaging tarkoittaa käyttäjien välisten viestien lähetystä lähes reaaliaikaisesti. Se eroaa sähköpostista käyttötavaltaan: viestit ovat yleensä lyhyitä ja niitä lähetetään nopeaan tahtiin. Suoraviestintä muistuttaa hyvin paljon keskustelua kahden tai useamman käyttäjän välillä. Varsinkin nuorten keskuudessa se on saavuttanut suuren suosion. Se on saamassa suosiota myös yrityskäytössä, koska se tarjoaa työntekijöille lisäkeinon kommunikoida työtovereiden ja asiakkaiden kanssa. Se on nopea ja kätevä vaihtoehto soittamiselle tai sähköpostin kirjoittamiselle.

Läsnäolo (presence) kertoo käyttäjän halukkuuden ja mahdollisuuden kommunikoida. Käyttäjä voi läsnäolotiedon avulla ilmaista, onko hän tavoitettavissa, ja jos niin mistä. Läsnäolotieto ei ole pelkkä online- tai offline-tieto vaan se kertoo enemmän. Esimerkiksi käyttäjä voi olla online, mutta varattu puhelinkeskustelun takia. Läsnäolotieto yhdistetään usein suoraviestintäsovellukseen, koska suoraviestintä vaatii kahden käyttäjän olevan halukkaita viestimään toisilleen. Itse suoraviestin lähettämiseen riittää kahden käyttäjän oleminen online-tilassa.

Suoraviestinnällä ja läsnäololla ei ole olemassa yhtenäistä standardia, vaan kaikki ohjelmistot käyttävät omia toteutustapojaan. Tämän vuoksi niiden ei ole mahdollista toimia keskenään. SIP yrittää olla standardi, jonka avulla eri ohjelmistojen olisi mahdollista toimia keskenään. Näin käyttäjät voisivat vapaasti valita haluamansa ohjelmiston, eikä sen takia että sitä on ”pakko” käyttää. Tällä hetkellä käyttäjä saattaa joutua käyttämään useaa eri suoraviestintäsovellusta halutessaan kommunikoida kaikkien kontaktiensa kanssa, koska osa niistä käyttää esimerkiksi AOL Instant Messenger-, osa Yahoo! Messenger- sekä loput MSN Messenger –sovellusta.

Suoraviestinnän merkitys yritysmaailmassa on jatkuvasti kasvamassa. Se tarjoaa yhden lisäkeinon kommunikoida työtovereiden tai asiakkaiden kanssa. Sen nopeus ja matala käyttökynnys lisäävät sen kiinnostavuutta. Matkapuhelinten yleistyessä ihmiset ovat nykyään näennäisesti tavoitettavissa. Näennäisesti siksi, että heidän oletetaan kuljettavan matkapuhelinta mukanaan. Läsnäolotiedon myötä käyttäjät saavat varmuuden toisen osapuolen tavoitettavuudesta.

SIP on suhteellisen uusi protokolla, vaikkakin se pohjautuu hyvin pitkälti HTTP:hen. Sen suurimmat vahvuudet ovat, että se on suunniteltu toimimaan yhdessä muiden web-teknologioiden kanssa sekä että siihen voi luoda laajennuksia. Arkkitehtuuri on suunniteltu siten, että mikäli jotain SIP-viestiä ei tunnisteta se välitetään eteenpäin muuttumattomana. Tämä mahdollistaa uusien toimintojen lisäämisen päätelaitteisiin ilman verkkoelementtien päivittämistä. Näin uusien toimintojen markkinoilletuloaika lyhenee huomattavasti.

SIP ei kuitenkaan ole täysin ongelmaton protokolla. Se tarjoaa tietoturvamekanismeja, mutta koko SIP-viesti on hankala salata. Verkkoelementtien pitää pystyä lukemaan osa viestistä, jotta ne osaavat välittää niitä. Varsinkin langattomissa päätelaitteissa ja verkoissa kaistanleveys on tärkeä kriteeri. SIP on tekstipohjainen, joten se ei ole kaikkein tehokkain tiedonsiirtotapa. Siihen kehitettyjen laajennuksien avulla viestejä voidaan pakata ja näin saada tiedonsiirto tehokkaammaksi. Suurin uhkakuva SIP:ille kuitenkin on kilpaileva protokolla. Jabber on vastaava protokolla kuin SIP:kin. SIP:in standardointityö on edennyt pidemmälle, mutta Jabberia käytetään enemmän. On vielä liian aikaista sanoa kumpi, jos kumpikaan, kilpajuoksun tulee voittamaan.

Tutkielmassani pyrin selvittämään, miten hyvin SIP soveltuu suoraviestinnän sekä läsnäolon viestintämekanismiksi, soveltuuko SIP tähän tarkoitukseen ja mitä mahdollisia ongelmia sen käyttöön saattaa liittyä. Lisäksi pyrin selvittämään, miten suoraviestintä ja läsnäolopalvelu toimivat. Pyrin kiinnittämään myös erityistä huomiota tietoturvaan, koska SIP toimii Internetissä, jota voidaan pitää vihamielisenä ympäristönä.

Tutkielman rakenne on seuraava: Luvussa 2 selitän yleisesti Internetin rakenteen, sen toimintaperiaatteen ja miten SIP sijoittuu suhteessa Internetiin. Luvussa 3 esittelen SIP:in ja kerron sen toimintaperiaatteen, arkkitehtuurin ja esittelen vastaavia protokollia. Luku 4 keskittyy suoraviestinnän esittelemiseen ja suoraviestinnän mahdollisuuteen SIP:issä. Luvussa 5 esittelen läsnäolon yleisen toimintaperiaatteen ja SIP:in mekanismit läsnäolon toteuttamiseksi. Luvussa 6 käsitellään SIP:in tietoturvaa, erilaisia mekanismeja sen saavuttamiseksi ja mahdollisia tietoturvauhkia. Lopuksi luvussa 7 pohdin SIP:in soveltuvuutta suoraviestintään, läsnäolon esittämiseen sekä SIP:in tulevaisuutta.

Tässä tutkielmassa eri verkkoelementtejä käsitellään selvyyden vuoksi erillisinä palvelimina. Läsnäolotietopalvelua kuitenkin käsitellään yhdistettynä paikkatietopalveluun ja niiden oletetaan sijaitsevan samassa sovellus-palvelimessa.

2. Internet

Internetin tarkoitus on tietoliikenneyhteyksien tarjoaminen. Muilla verkoilla on muita tarkoituksia, esimerkiksi televerkon tarkoituksena on telepalveluiden tarjoaminen ja tv-verkon tarkoituksena on tv-ohjelmien lähettäminen. Internet koostuu useista pienistä verkoista, jotka ovat yhteydessä toisiinsa.

Internetin vahvuuksia ovat sen vikasietoisuus sekä soveltuvuus palveluiden luontialustaksi. Vikasietoisuutta lisää, ettei verkko tarvitse tilatietoa lähettääkseen viestejä eli ip-paketteja vastaanottajalle. Tällöin verkon solmun eli reitittimen lakatessa toimimasta liikenne ohjataan perille toisen solmun kautta. Vikasietoisuutta lisää myös, että verkko koostuu älykkäistä pääte-laitteista. Itse reitittimet ovat passiivia ja lähettävät ip-paketteja saamiensa ohjeiden mukaisesti. Internetin koostuminen useista eri kerroksista helpottaa sovellusten rakentamista. Sovelluksen kehittäjä valitsee tarvittavat protokollat, jotka huolehtivat datan siirrosta.

Internet protokollapino koostuu useista kerroksista, joita ovat (kuva 1) peruskerros (physical layer), siirtokerros (data link layer), verkkokerros (network layer), kuljetuskerros (transport layer) ja sovelluskerros (application layer). Kuvassa 1 on myös esitetty jokaisen kerroksen siirrettävän datan nimitys eli PDU (protocol data unit).

Peruskerros koostuu fyysisistä laitteista. Siirtokerros määrittelee miten yhteyksiä rakennetaan solmujen välille. Verkkokerroksen, eli IP-kerroksen, tehtävänä on reitittää lähetettävät paketit lähettävältä isäntäkoneelta vastaanottavalle isäntäkoneelle useista erityyppisistä aliverkoista koostuvan yhteyden yli. Kuljetuskerroksen tehtävänä on siirtää sovelluskerroksen sanomat asiakkaalta palvelimelle ja päinvastoin. Sovelluskerros on tarkoitettu sovelluksen eri komponenttien väliseen viestintään.

Kuva 1. Internetin protokollapinon kerrokset

Siirtokerros määrittelee miten yhteyksiä rakennetaan verkossa solmusta toiseen ja miten tieto kulkee luotettavasti siirtotietä pitkin. Tämän kerroksen protokollat huolehtivat virhesuojauksesta peruskerroksessa, tietovuon kontrol-loinnista ja kehysten synkronisoinnista.

Verkkokerroksen tehtävänä on reitittää datagrammit lähettävältä isäntäkoneelta vastaanottavalle isäntäkoneelle. Internet käyttää tähän Internet-protokollaa (IP). IP mahdollistaa verkon liittämisen muihin verkkoihin riippumatta niiden teknologisista ratkaisuista ja pakettien lähettämisen näiden verkkojen välillä. IP:tä käyttävät verkkokerroksessa reitittimet ja palvelimet siirtäen näin älykkyyttä päätelaitteille. Päätelaitteet ovat vastuussa ip-pakettien liikenteen kontrolloinnista, protokolla ei varmista pakettien perillemenoa eikä pakettien järjestystä [RFC 791]

Kuljetuskerroksen protokollien perustehtävänä on tarjota sovelluksille kuljetuspalvelua kahden mahdollisesti eri koneissa tai jopa eri verkoissa olevan prosessin välillä. Kuljetusprotokollan vähimmäisvaatimuksena on kyky osoittaa lopullinen kohde eli se sovellus tai prosessi, jonka kanssa halutaan kommunikoida. TCP/IP-arkkitehtuurissa on kaksi vaihtoehtoa kuljetus-protokollaksi: TCP ja UDP.

Transmission Control Protocol eli TCP tarjoaa luotettavan yhteydellisen kuljetuspalvelun, joka takaa pakettien järjestyksen säilymisen sekä pakettien kuljetuksen vastaanottajalle. Tämän ansiosta sovellusohjelmoijan ei tarvitse kiinnittää huomiota tiedonsiirrossa esiintyvien ongelmien hoitamiseen vaan TCP hoitaa ne hänen puolestaan. TCP perustuu lähetysikkunan käyttöön, jossa sovellus lähettää useampia paketteja kerrallaan ennen kuin jää odottamaan kuittausta. Vastaanottajan puolestaan ei tarvitse lähettää kuittausta kaikista vastaanottamista paketeista. Vastaanottaja kuittaa kertomalla kuinka pitkälle se on vastaanottanut dataa oikein. TCP:ssä ei kuitenkaan kuitata paketteja

niiden järjestysnumeron perusteella vaan alkuperäisestä datasta vastaanotettujen tavujen perusteella. Käytännössä tämä tarkoittaa sitä, että kuittauksessa kerrotaan mitä tavua vastaanottaja seuraavaksi odottaa lähettäjältä. Mikäli vastaanottaja ei saa jotain pakettia tai sen sisältö on muuttunut virheelliseksi siirron aikana, vastaanottaja ei lähetä kuittausta. Tietyn odotusajan jälkeen lähettäjän odotusaikalaskuri nollautuu ja kaikki data viimeisessä kuittaus-sanomassa olleesta tavusta alkaen lähetetään uudelleen [RFC 793].

Kaikkien pakettien vastaanottaja on sama eli koneen ip-osoite. Paketti kohdistetaan oikealle sovellukselle TCP-porttinumeroinnin avulla, esimerkiksi porttiin 80 tulevat paketit lähetetään HTTP-sovellukselle.

User Datagram Protocol eli UDP tarjoaa yksinkertaisen yhteydettömän kuljetuspalvelun sovelluksille. Käytännössä se ei lisää alemman protokollan eli IP:n päälle muuta kuin mekanismin osoittaa lopullinen kohde. UDP ei sisällä pakettien numerointia, eikä siten anna varmuutta pakettien järjestyksen säilymisestä. Koska lähetettäviä paketteja ei numeroida, vastaanottaja ei myöskään voi kuitata niitä vastaanotetuiksi. Tämän vuoksi UDP:tä voidaan pitää epäluotettavana kuljetuspalveluna. UDP:ssä virheiden havaitsemiseksi lasketaan tarkistussumma koko UDP-paketista eli myös datakentästä. UDP-tarkistussumman käyttö on ainoa tapa taata datan oikeellisuus siirron jälkeen. Kuten TCP:ssäkin kaikkien pakettien vastaanottaja on sama eli koneen ip-osoite. Paketti kohdistetaan oikealle sovellukselle porttinumeroinnin avulla [RFC 768].

Internet-arkkitehtuurin ylimmällä tasolla ovat varsinaiset sovellukset. Sovelluskerroksen tarkoituksena on välittää eri sovellusten viestejä toisille sovelluksille. Sovelluskerros tarjoaa lukuisia eri protokollia, joista sovellus valitsee sen tarvitsemat protokollat. Esimerkkejä näistä protokollista ovat Hypertext Transfer Protocol eli HTTP, File Transfer Protocol eli FTP, Simple Mail Transfer Protocol eli SMTP sekä Session Initiation Protocol eli SIP.

3. SIP

3.1. SIP:in historia

SIP:in kehittely alkoi 1990-luvun puolivälissä Columbian yliopistossa Henning Schulzrinnen johdolla. Tarkoituksena oli luoda standardi, Multiparty Multimedia Session Control – MMUSIC, jolla voitaisiin kontrolloida audio-visuaalista dataa tietoverkoissa. Tämän standardin pohjalta luotiin ensimmäiset versiot SIP:istä.

Ensimmäinen versio SIP:stä (SIPv1) ilmestyi helmikuussa 1996 ja oli nimeltään Session Invitation Protocol. Sen kehittivät Mark Handley ja Eve Schooler ja sen julkaisi IETF-statuksella (Internet Engineering Task Force) ”Internet draft”. SIPv1 käytti Session Description –protokollaa (SDP) istuntojen kuvaamiseen ja UDP:tä tiedon siirtämiseen. SIPv1 oli tekstipohjainen ja käsitteli pelkästään istuntojen perustamista. Merkinanto loppui käyttäjän liittyessä istuntoon. Istunnon aikaisia kontrolleja ei SIPv1:ssä myöskään ollut [Camarillo].

Simple Conference Invitation –protokollan (SCIP) julkaisi Henning Schulzrinne myöskin helmikuussa 1996 IETF-statuksella ”Internet draft”. Myös SCIP oli mekanismi käyttäjien kutsumiseksi kaksipisteyhteysistuntoihin (point-to-point). Se perustui HTTP:hen ja käytti siten TCP:tä kuljetusprotokollanaan. SIPv1:n lailla se oli tekstipohjainen. SCIP käytti käyttäjien sähköpostiosoitteita tunnuksina, joiden tarkoituksena oli tarjota universaali käytäjätunnus [Camarillo].

Session Initiation Protocol (SIPv2) syntyi SIPv1:n ja SCIP:n yhdistämisen tuloksena. SIPv1 ja SCIP olivat varsin samankaltaiset protokollat, eikä ollut järkevää jatkaa molempien kehitystä. Lopputulos piti lyhenteen SIP, mutta vaihtoi lyhenteen merkitystä. SIPv2:n Internet draftin kirjoittivat Hanley, Schulzrinne ja Schooler ja se julkaistiin joulukuussa 1996. Uusi SIP yhdisti vanhojen protokollien parhaat puolet, se perustui HTTP:hen, mutta pystyi käyttämään kuljetusprotokollanaan sekä TCP:tä että UDP:tä. SDP:tä käytettiin kuvaamaan istuntoja; SIPv2 oli tekstipohjainen.

IEFT perusti SIP-työryhmän (SIP Working Group) syyskuussa 1999. Maaliskuussa 2001 työryhmä jaettiin kahtia. Ensimmäinen työryhmä keskittyy itse SIP:in määrittelyyn ja sen lisäosiin, vastaavasti jälkimmäinen, SIPPING-työryhmä, keskittyy SIP:iä käyttäviin sovelluksiin [Camarillo].

”3rd Generation Partnership Project” (3GPP) on eri telekommuni-kaatiostandardointijärjestöjen välinen yhteistyöelin, jonka tavoitteena on luoda maailmanlaajuinen standardi kolmannen sukupolven (3G) televerkoille.

Vuonna 2000 se valitsi SIP:in protokollaksi luotaessa multimediaistuntoja 3G-verkoissa.

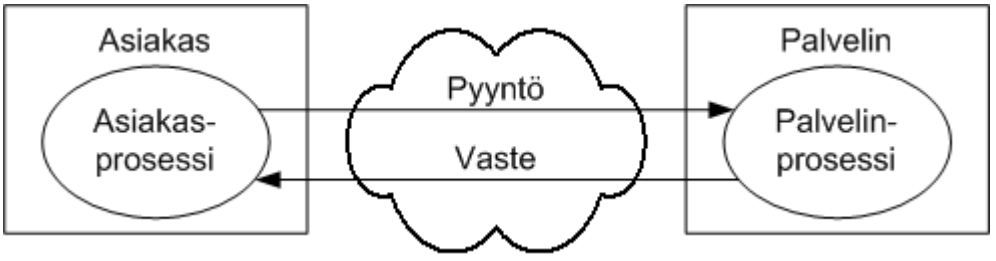
SIP ei ole vanha protokolla, itse asiassa viimeisin standardi, RFC 3261, saatiin valmiiksi kesäkuussa 2002. Uutuutensa vuoksi markkinoilla on olemassa useita SIP-toteutuksia, laitteita ja ohjelmistoja, jotka eivät vastaa standardia. On kaikkien etu saada toteutukset toimimaan standardin mukaisesti, jotta ne voivat keskustella keskenään ilman virhetilanteita. Lisäksi SIP:iä kehitetään jatkuvasti ja siihen luodaan uusia laajennoksia. Tämän vuoksi SIP-yhteisö on perustanut SIPit-tapahtuman (SIP interoperatibility test event).

SIPit-tapahtumia järjestetään yleensä kaksi tai kolme vuodessa jonkin messun tai muun tapahtuman yhteydessä. Näiden kaikille avointen tapahtumien tarkoituksena on testata eri SIP-toteutusten yhteensopivuutta, havaita virheitä toteutuksissä sekä tarvittaessa kommentoida ja ehdottaa parannuksia SIP-määrittelyihin. Tapahtumien tulokset ovat luottamuksellisia. Osallistujien ei ole lupa kommentoida, julkistaa tai muuten luonnehtia muiden osallistujien testien tuloksia.

3.2. SIP:in tarkoitus

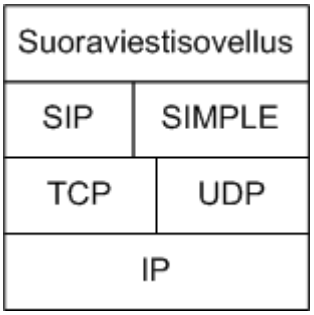
SIP luo, muokkaa ja päättää istuntoja. Sitä voidaan käyttää kutsumaan uusia käyttäjiä olemassa olevaan istuntoon tai luomaan uusi istunto. SIP ei ota kantaa itse istunnon sisältöön, vaan se voi olla mitä tahansa - esimerkiksi videokonferenssi, puhelu, jaettu työpöytä tai peli-istunto. Istunnot kuvataan yleensä SDP:llä, mutta kuvaamiseen voidaan käyttää myös muita kuvaus-protokollia [RFC 3261].

SIP perustuu HTTP:hen ja kuten HTTP:kin se noudattaa asiakas-palvelin -mallia (kuva 2). Asiakaskoneella oleva prosessi lähettää ip-paketteja suoraan palvelinkoneella olevalle prosessille, joka lähettää asiakkaalle vasteen [RFC 2616].



Kuva 2. Asiakas-palvelin –malli

SIP on sovelluskerroksen protokolla. Kuljetuskerroksen protokollista SIP voi käyttää sekä TCP:tä että UDP:tä. Verkkokerroksessa SIP kuten muutkin Internetin sovellukset käyttävät IP:tä. Kuvassa 3 on esitetty SIP:iä ja SIMPLE:ä käyttävän suoraviestisovelluksen protokollapino.



Kuva 3. Suoraviestisovelluksen protokollapino

3.3. SIP-verkon arkkitehtuuri

SIP-arkkitehtuuri käsittää kaksi peruskomponenttia, SIP-käyttäjäagentin (user agent – UA) ja SIP-palvelimen (network server) [Sinnreich and Johnston].

Käyttäjäagentti on loppukäyttäjälle tarkoitettu komponentti SIP-toimintojen suorittamiseksi. Se koostuu kahdesta eri osasta, asiakaselementistä (user agent client – UAC) ja palvelinelementistä (user agent server - UAS). Asiakas-elementin tehtävä on muodostaa yhteyksiä ja palvelinelementin tehtävä on vastaanottaa ja lähettää SIP-viestejä. Yleensä nämä kaksi elementtiä näkyvät loppukäyttäjälle yhtenä päätepisteenä (endpoint). Päätepiste voi olla hyvinkin kevyt ja tarkoitettu mobiileihin päätelaitteisiin tai se voi olla sulautettuna muihin sovelluksiin, esimerkiksi SIP-puhelin tai SIP-suoraviestintäohjelma. Käyttäjäagentin rooli vaihtelee tilanteen mukaan: esimerkiksi luodessaan yhteyttä se toimii asiakaselementtinä ja lähettäessään INVITE-viestiä se toimii palvelinelementtinä. Vain käyttäjäagentit pystyvät luomaan SIP-viestejä. SIP-palvelimetkin sisältävät käyttäjäagentteja, näin niiden on mahdollista lähettää viestejä [Sinnreich and Johnston].

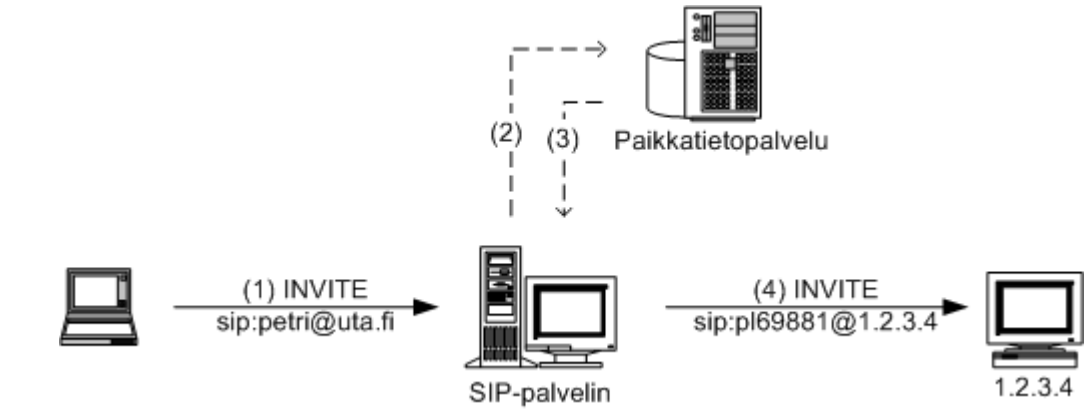
SIP-palvelimen päätehtävä on rekisteröidä, välittää ja hallinnoida SIP-viestien mukaisia toimintoja ja toimenpiteitä. SIP-palvelimia on kolmen tyyppisiä. Ne ovat joko välipalvelimia (proxy server), uudelleen-ohjauspalvelimia (redirect server) tai rekisteröintipalvelimia (registrar server).

Välipalvelimet vastaanottavat SIP-viestejä ja kyselevät paikkatieto-palvelulta (location service) vastaanottajien osoitetietoja. Osoitetiedon kyselyn jälkeen ne lähettävät SIP-viestin eteenpäin suoraan käyttäjälle, mikäli tämä sijaitsee samalla toimialueella (domain), tai toiselle välipalvelimelle, mikäli käyttäjä on toisella toimialueella. Välipalvelimet voivat lisätä viesteihin parametrejä tai kieltäytyä vastaanottamasta viestejä, mutta ne eivät voi luoda viestejä eivätkä vastata viesteihin myöntävästi. Mikäli ne eivät tunnista viestiä, ne välittävät sen eteenpäin muuttumattomana. Tämä mahdollistaa uusien

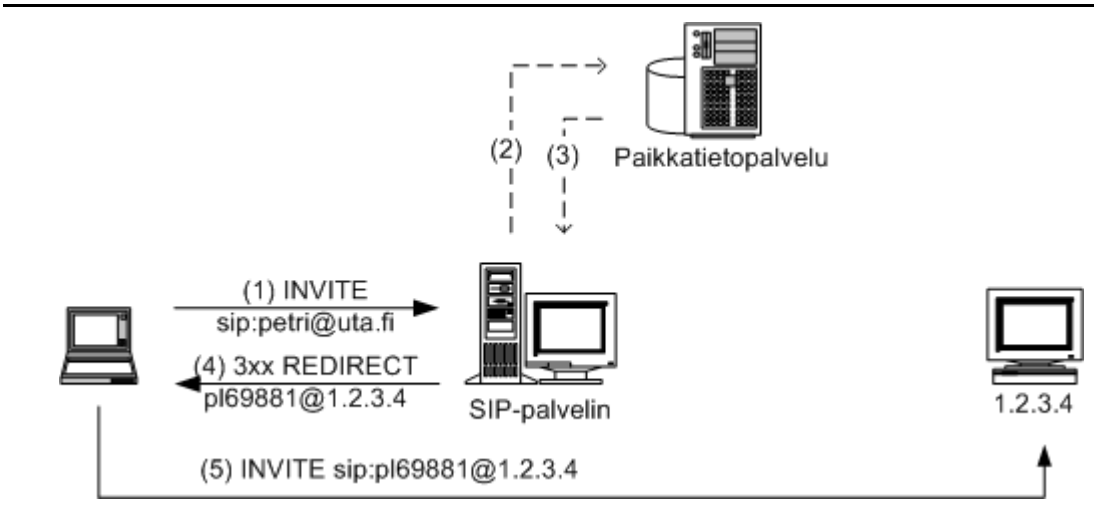
toimintojen lisäämisen suoraan käyttäjäagentteihin, eikä välipalvelinten ohjelmistoja tarvitse päivittää [RFC 3261].

Uudelleenohjauspalvelimet vastaanottavat SIP-viestejä ja kyselevät paikkatietopalvelulta vastaanottajien osoitetietoja. Tämän jälkeen ne lähettävät käyttäjälle 3xx-tyyppisen viestin ohjaten käyttäjän luomaan yhteyttä toisesta ip-osoitteesta. Ne eivät lähetä viestiä eteenpäin sen vastaanottajalle. Päätelaitteen vastuulle jää viestin lähettäminen eteenpäin palvelimelta saamiensa tietojen pohjalta [RFC 3261].

Näiden kahden eri palvelintyypin toiminnallista eroa on havainnollistettu kuvissa 4 ja 5. Kuvassa 4 oleva SIP-palvelin toimii välipalvelin periaatteella ja kuvassa 5 oleva SIP-palvelin toimii uudelleenohjausperiaatteella. Vastaanottaessaan SIP-viestin välipalvelin kyselee paikkatietopalvelulta mistä viestin vastaanottaja löytyy ja välittää viestin eteenpäin vastaanottajalle. Vastaavasti uudelleenohjauspalvelin kyseltyään käyttäjän ip-osoitteen lähettää viestin takaisin päätelaitteelle, joka lähettää viestin suoraan vastaanottajalle. Kuvissa oletetaan, että kaikki laitteet sijaitsevat samalla toimialueella, eikä viestejä lähetetä toisille toimialueille.

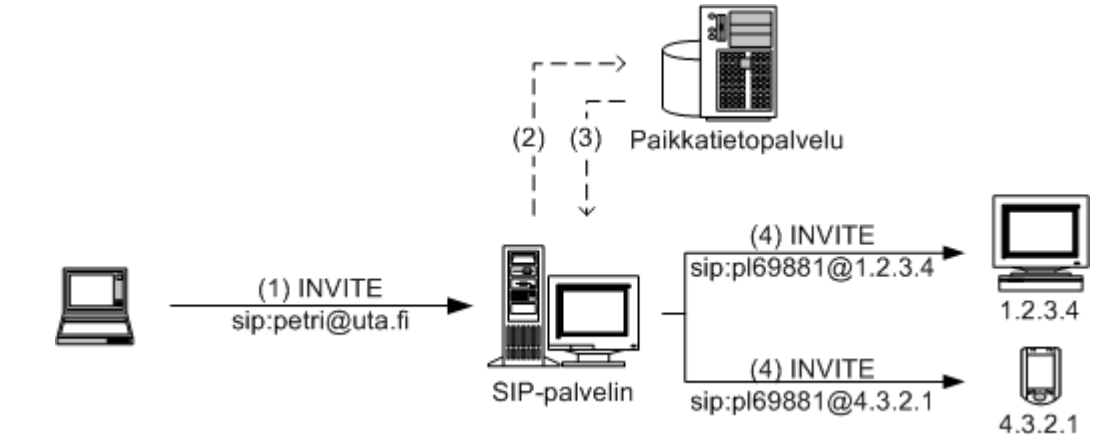


Kuva 4. SIP-välipalvelimen toimintaperiaate



Kuva 5. SIP-uudelleenohjauspalvelimen toimintaperiaate

SIP-palvelin pystyy myös jakamaan (forking) saamansa viestin useaan eri osoitteeseen (kuva 6). Käyttäjän rekisteröityä useita eri päätelaitteita, ne kaikki hälyttävät tulevasta viestistä. Viestiin voidaan vastata mistä tahansa päätelaitteesta. Kuvassa 6 oleva välipalvelin kyselee paikkatietopalvelulta mistä käyttäjä löytyy. Tämän jälkeen se lähettää saamansa viestin molempiin käyttäjän rekisteröimiin päätelaitteisiin.



Kuva 6. SIP-viestin jakaminen

Rekisteröintipalvelimet ovat palvelimia, jotka vastaanottavat käyttäjäagenttien lähettämiä paikkatietoviestejä, käytännössä ip-osoitteita. Vastaanotet-tuaan viestin ne pyytävät paikkatietopalvelua varastoimaan käyttäjän paikka-tiedon. Usein paikkatietopalvelu on toteutettu fyysisesti rekisteröinti-palvelimeen. Näin ei kuitenkaan tarvitse olla, vaan paikkatietopalvelu voi sijaita myös sovelluspalvelimessa.

Paikkatietopalvelu on palvelu, johon rekisteröintipalvelin tallentaa saamiaan REGISTER-viestien tietoja. Välipalvelimet kyselevät paikkatietopalvelulta saamiensa viestien vastaanottajien ip-osoitteita [RFC 3261].

Sovelluspalvelimet (application servers) toimivat yhdessä SIP-palvelinten sekä käyttäjäagenttien kanssa tarjoten näille palveluita. Tälläisiä ovat muun muuassa suoraviestintä, läsnäolo ja käyttäjän profilointi.

AAA-palvelin (Authentication, Authorization and Accounting) tarjoaa todennus-, valtuutus- sekä hallinnointitoimintoja. Nämä tehtävät voidaan jakaa myös useampien AAA-palvelinten kesken, jolloin käyttäjän todentava palvelin ei välttämättä ole palvelin, joka valtuuttaa toisia käyttäjiä käyttämään palveluita.

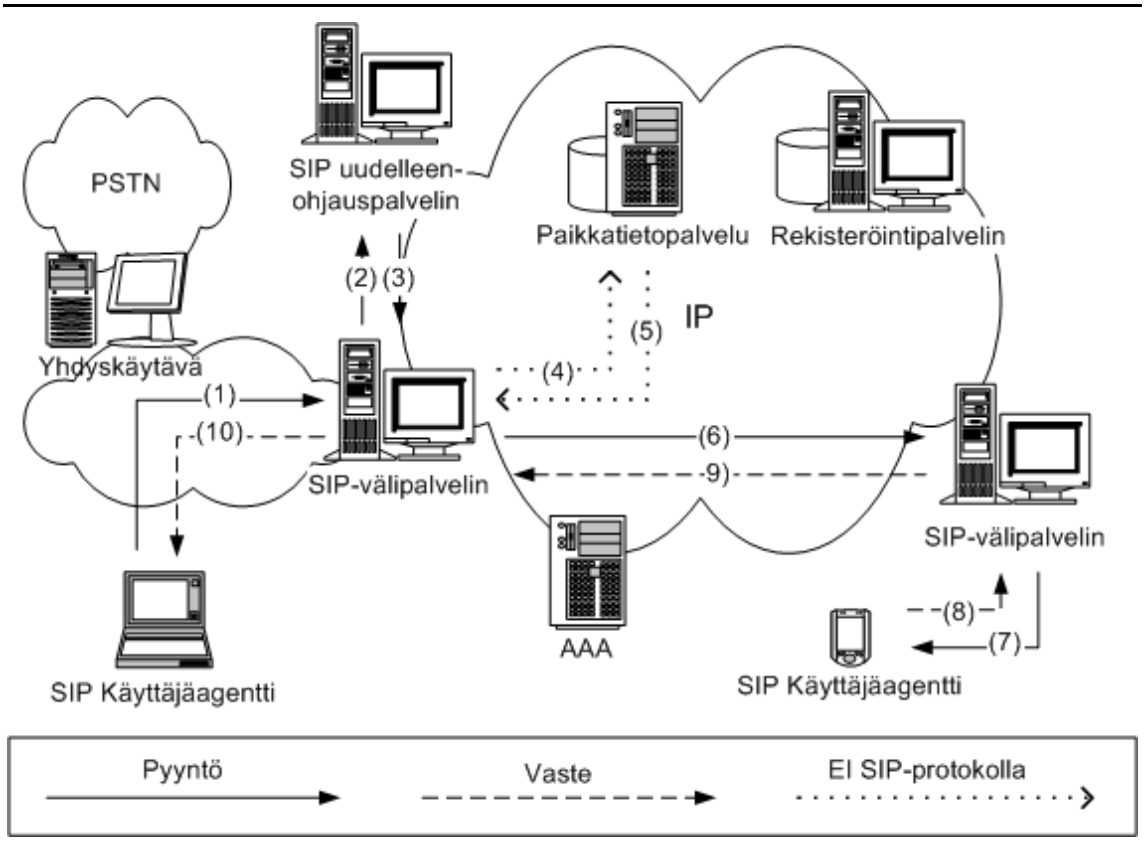
Back-to-Back-käyttäjäagentti (B2BUA) on palvelu, joka vastaanottaa ja prosessoi viestejä kuin se olisi käyttäjäagentin palvelinelementti. Ymmärtääkseen miten viesteihin tulisi vastata se toimii kuten käyttäjäagentin asiakaselementti. Toisin kuin välipalvelin, se ylläpitää tilatietoa keskustelusta (dialog) ja vastaa kaikkiin sen keskustelun viesteihin. Tällä tavalla on mahdollista muun muassa toteuttaa anonyymipalvelu [RFC 3261].

SIP-yhdyskäytävää (gateway) käytetään kun halutaan toteuttaa jokin toiminto tai palvelu, joka sijaitsee muussa kuin IP-verkossa. Yhdyskäytävä voi olla tyypiltään esim:

- PSTN-yhdyskäytävä (perinteinen puhelinverkko),
- MOBILE-yhdyskäytävä (langaton puhelinverkko),
- PBX-yhdyskäytävä (sisäinen puhelinkytkentäverkko),
- Frame relay / ATM-yhdyskäytävä.

SIP-istunnon luonnin jälkeen SIP ei enää tarvitse välipalvelimia viestien välittämiseen. SIP-käyttäjäagentit lähettävät viestejä suoraan toisilleen (point-to-point). Käytännössä SIP-viestit usein kulkevat välipalvelinten kautta, varsinkin päätelaitteiden ollessa eri toimialueilla. Teoriassa päätelaitteet voivat kuitenkin lähettää viestejä suoraan toisilleen.

Kuvassa 7 on esitetty laajempi SIP-verkon arkkitehtuuri. Kuvassa on myös esitetty tavallisen SIP-yhteydenottopyynnön (INVITE-viesti) kulku verkossa. Kuvassa ei ole otettu huomioon käyttäjien tunnistusta eikä valtuutusta käyttää palveluita tai tietoverkkoa, vaikka kuvassa onkin AAA-palvelin.



Kuva 7. SIP-verkon arkkitehtuuri

SIP:illä on kaksi ongelmaa verkkoympäristöissä. Ensimmäinen ongelma ilmenee käytettäessä palomuureja, ja toinen ongelma koskee ip-osoitteiden muuntamista verkossa (Network Address Translation – NAT).

Palomuurit voidaan helposti määrittää sallimaan SIP-viestintä, koska SIP käyttää vakioporttia 5060. Ongelmaksi muodostuu itse istunnon liikenne. Istunto käyttää viestintään RFC 1887 –dokumentissa määriteltyä RTP:tä (Real-time Transport Protocol), joka koostuu data- ja kontrollointiosasta. RTP:n ongelma on, ettei se käytä tiettyjä portteja vaan portit määritellään dynaamisesti. Lisäksi se ei varsinaisesti ole oma protolla. Näiden seikkojen vuoksi useat palomuurit eivät osaa tunnistaa RTP-liikennettä ja näin päästää sitä läpi.

Toinen ongelma liittyy ip-osoitteiden muuntamiseen. NAT-muunnoksessa verkon sisäpuolelta tuleva ip-liikenne näyttää tulevan yhdestä ip-osoitteesta. Muunnoksessa vaihdetaan kaikkien ip-pakettien otsakekenttien ip-osoitteet toimialueen NAT-palvelimen ip-osoitteeksi. Menetelmällä pyritään suojaamaan verkon yksityisyyttä tai luomaan lisää ip-osoitteita verkon sisäpuolelle. Ongelma koskee SIP:iä, koska ip-osoitteita on myös pakettien sisällä eikä pelkästään ip-pakettien otsakekentissä. Näin ip-osoite paketin sisällä ei vastaa ip-osoitetta paketin ulkopuolella.

3.4. Vastaavat protokollat

SIP ei suinkaan ole ainutlaatuinen protokolla vaan on olemassa myös muita vastaavia protokollia. SIP:in suurimpana etuna on, että se on kehitetty toimimaan yhdessä web-teknologioiden kanssa eikä siitä ole yritetty luoda kaiken kattavaa ratkaisua.

H.323 ja MGCP / Megaco ovat samankaltaisia protokollia, joita voidaan käyttää toteuttamaan samoja toimintoja kuin SIP:kin. Jabber on suunniteltu samaan tehtävään kuin SIP:kin, joten niiden välillä on enemmän kilpailua standardin asemasta. Tällä hetkellä SIP on edennyt standardoinnissa pidemmälle, mutta Jabberia käytetään enemmän.

On myös olemassa lukuisia muita ohjelmistoja, kuten Yahoo! Messenger, ICQ, MSN Messenger ja AOL Instant Messenger, jotka käyttävät omia suljettuja protokollia viestien välitykseen. Tämä onkin ollut suoraviestinnän ongelma pitkän aikaa. Ei ole ollut mahdollista kommunikoida kahden eri ohjelman välillä. Näiden ohjelmien suuren käyttäjämäärän takia tilanne tuskin tulee lähitulevaisuudessa nopeasti muuttumaankaan.

3.4.1. H.323

H.323 on osa reaaliaikaista kommunikointiprotokollaperhettä. Se on ITU:n (Internation Telecommunication Union) hyväksymä standardi ja kuuluu H.32x standardiperheeseen. Kaikki standardiperheen standardit nojaavat vahvasti H.320 –standardiin Jokainen protokolla tässä perheessä on suunniteltu toimimaan eri verkkoympäristössä. H.323 määrittelee miten videokokousten data välitetään LAN-verkoissa.. H.323-arkkitehtuuri sisältää seuraavat elementit:

- Yhdyskäytävä (gateway), joka yhdistää LAN-verkon H.323-loppukäyttäjät muiden verkkojen loppukäyttäjiin. Ne muuntavat protokollia, konvertoivat mediaa toisiin formaatteihin ja siirtävät tietoa.
- ”Gatekeeper”:ien tehtävä on muuntaa osoitteita, varata kaistanleveyttä ja tarjota muita kontrollointi- ja hallintatoimintoja. Ne ovat H.323-verkon ”aivot” ja toimivat kun SIP-palvelimet.
- ”Multipoint control units” yhdistävät ja jakelevat konferenssin tietovirtaa kolmelle tai useammalle H.323 terminaalille.

SIP:in suurin ero H.323-protokollaan palveluiden näkökulmasta on, että SIP mahdollistaa suoraviestien lähetyksen ja läsnäolotiedon välityksen. H.323 ei näitä palveluita mahdollista.

Teknisestä näkökulmasta SIP eroaa H.323:sta modulaarisuutensa ja muutettavuutensa vuoksi. SIP on modulaarinen standardi, joka on suunniteltu

web-teknologioiden ympärille ja toimimaan niiden kanssa. H.323 vastaavasti on ”sateenvarjo”standardi, jolta puuttuu osa siitä joustavuudesta, jonka SIP-arkkitehtuuri tarjoaa. SIP ei luo tilanteita, joissa jokainen uusi laajennos tarvitsee oman ohjelmansa. Lisäksi kaikki palveluntarjoajat voivat olla vastuussa omista palveluistaan, koska ohjelmat muodostuvat selkokielisestä tekstistä. H.323:ssa vastaavasti koodi on binäärimuodossa. Samalla palvelun-tarjoajat tulevat myöskin vähemmän riippuvaisiksi laitevalmistajista.

3.4.2. MGCP/Megaco

MGCP / Megaco (Media Gateway Control Protocol / Media Gateway Controller) on telekommunikaatioyhteisön kehittämä standardi, jonka tarkoituksena on yhdistää SS7 (Signalling System 7) ja VoIP (Voice over IP). Sen tarkoitus oli korvata H.323-standardi, joka oli kasvanut liian suureksi eikä ollut yhteen-sopiva puhelinverkkojen kanssa. Tämän ongelman ratkaisemiseksi H.323-mallin mukainen signalointi poistettiin yhdyskäytävästä, ja sen tilalle luotiin ”media gateway controller” tai ”softswitch”. Tämä laite pystyy kontrolloimaan useita ”media gateway”-laitteita.

MGCP / Megaco on isäntä-orja –kontrollointiprotokolla (master / slave control protocol) jota käytetään ”softswitch”- ja ”media gateway”-laitteiden väliseen kommunikointiin. ”Media Gateway Controller” tai ”softswitch” on aina isäntä ja ”media gateway” orja. Isäntä määrää kaikki toiminnot ja orja pelkästään suorittaa tehtäviä. SIP:issä vastaavasti jokainen asiakas voi luoda yhteyden toiseen asiakkaaseen.

Suurin ero SIP:iin nähden on, ettei MGCP / Megacoa ole tarkoitettu koko järjestelmän protokollaksi. Se tarvitsee SIP:iä kommunikoimaan ”softswitch”-laitteiden välillä. SIP:in samankaltaisuus muiden web-teknologioiden kanssa suosii sen käyttöä IP-verkoissa.

3.4.3. Jabber

Jabber on kokoelma standardeja, jotka mahdollistavat kahden käyttäjän välisten viestien, läsnäolo-tiedon tai muun järjestetyn tiedon välittämisen Internetissä lähes reaali-aikaisesti. Se on täysin XML-pohjainen, ja sitä kehittää IETF:n alaisuudessa oleva ”Extensible Messaging and Presence Protocol”-työryhmä (xmpp) avoimen lähdekoodin periaatteella. Se on hyvin samankaltainen kuin SIP ja sen tavoitteetkin ovat samat kuin SIP:in.

On huomioitava että markkinoilla on olemassa myös maksullinen Jabber-niminen ohjelmisto, jota kehittää yritys nimeltä Jabber (www.jabber.com). Vastaavasti Jabber-yhteisön ylläpitämä sivusto on osoitteessa www.jabber.org.

Jabber-nimen käyttö saattaa aiheuttaa sekaannuksia, tässä tutkielmassa Jabber-nimellä viitataan aina standardiin ellei toisin mainita.

Jabber käyttää asiakas-palvelin -arkkitehtuuria, jossa asiakas ottaa TCP-yhteyden palvelimeen ja keskustelee XML-viestein. Kaiken Jabber-liikenteen pitää kulkea vähintään yhden Jabber-palvelimen kautta. Toisin kuin SIP, se ei salli suoraa liikennettä kahden asiakkaan välillä. Asiakkaan avatessa yhteyden palvelimeen, Jabber avaa yhdensuuntaisen XML-tietovirran asiakkaalta palvelimelle. Vastaavasti palvelin vastaa avaamalla samanlaisen tietovirran palvelimelta asiakkaalle. Kaikki viestintä Jabberissa tapahtuu näitä tietovirtoja käyttäen. Tietovirrat ovat olemassa, kunnes toinen osapuoli lopettaa istunnon. Itse viestien välitys ja muut toiminnot ovat Jabberissa hyvin paljon samankaltaisia kuin SIP:issäkin. Usein vain niiden toteutustapa poikkeaa toisistaan.

Jabber-palvelimilla on pääasiassa kolme eri tehtävää: ne hallinnoivat ja kommunikoivat suoraan Jabber-asiakkaiden kanssa, kommunikoivat toisten Jabber-palvelinten kanssa sekä hallinnoivat useita palvelinkomponetteja (kuten rekisteröinti, autentikointi ja läsnäolo). Jabber-asiakkaiden ainoat tehtävät ovat kommunikoida palvelinten kanssa käyttäen TCP-yhteyttä sekä tulkita ja ymmärtää vastaanottamiaan XML-viestejä.

Jabberin tiedonsiirtoprotokolla on XMPP. Se täyttää RFC 2778 –määritelmässä esitetyt vaatimukset suoraviestinnälle ja läsnäololle, muttei ole yhteensopiva SIP:in kanssa. Peruselementtejä siinä ovat </message>, </presence> ja </iq>. Message-elementtiä käytetään suoraviestien välitykseen (esimerkki message-elementistä on esitetty koodiesimerkissä 1). Presence-elementtiä käytetään läsnäolotiedon välittämiseen. Iq-elementtiä (Info/Query) käytetään erilaisten tiedontojen sekä kyselyiden välittämiseen.

```
<message from='petri@uta.fi' to='lintula@cs.uta.fi'>
  <body>Hello world!</body>
</message>
```

Koodiesimerkki 1. Jabber message-viesti

Jabber on SIP:iä kevyempi. Sen viesteissä välitetään vähän tietoa, joka havaitaan koodiesimerkeistä 1 ja 2. On kuitenkin huomioitava, että koodiesimerkeissä esitetyt viestit ovat hyvin pelkistettyjä ja ne voivat sisältää enemmänkin kenttiä. Viestien lyhyys on etu, koska varsinkin langattomissa verkoissa välitettävän datan määrä tulee pitää minimissään.

```
<presence>
```

```
  <show>dnd</show>
  <status>lunch</status>
</presence>
```

Koodiesimerkki 2. Jabber presence-viesti

Standardoinnissa SIP on edennyt pidemmälle kuin Jabber. SIP:in laajennuksen suoraviestintää varten eli SIMPLE:n standardointityö ei kuitenkaan ole yhtä pitkällä kuin SIP:in standardointityö. Jabberia käyttäviä sovelluksia on olemassa enemmän kuin SIP:iä käyttäviä sovelluksia. Tämä saattaa muuttua tulevaisuudessa, koska Microsoft sekä IBM ovat päättäneet tukea SIP:iä.

Markkinoilla on tällä hetkellä useita suoraviestintä ohjelmistoja jotka käyttävät niiden omia suljettuja ratkaisujaan, kuten MSN Messenger, AOL Instant Messanger, ICQ sekä Yahoo! Messenger. Voidaankin todeta, ettei kumpikaan SIP eikä Jabber johda kilpailua avoimesta suoraviestintä- ja läsnäolostandardista.

3.5. SIP-käyttäjät

SIP ei voi kutsua käyttäjää istuntoon, ellei se tiedä mistä hänet tavoittaa. Käyttäjää ei ole sidottu tiettyyn paikkaan vaan hän on tavoitettavissa eri ip-osoitteista. Käyttäjä voi ohjata yhteyspynnöt myös eri päätelaitteille: aamulla hän ohjaa saamansa yhteyspyynnöt tietokoneelleen, iltapäivällä pda-laitteelleen ja illalla matkapuhelimeensa [RFC 3261].

Käyttäjä tunnistetaan SIP-ympäristössä SIP Internet-resurssin tunnisteen (Uniform Resource Indicator - URI) avulla. SIP Internet-resurssin tunnisteen muoto on samanlainen kuin sähköpostiosoitteenkin. Se koostuu käyttäjänimestä ja verkkotunnuksesta, esimerkiksi ”sip:petri.lintula@uta.fi”. SIPS Internet-resurssin tunniste on vastaava kuin SIP Internet-resurssin tunnistekin, mutta siinä yhteys luodaa suojatusti (securely). Mikä tahansa SIP Internet-resurssin tunniste voidaan muuttaa SIPS Internet-resurssin tunnisteeksi vaihtamalle skeema (scheme). Näin voidaan aina kommunikoida käyttäen suojattua yhteyttä. Kuten kaikkia Internet-resurssin tunnisteita, SIP- ja SIPS-osoitteita voidaan käyttää web-sivuilla, sähköposteissa tai tulostettuina. Ne sisältävät riittävästi tietoa, jotta istunto halutun vastaanottajan kanssa voidaan luoda ja ylläpitää.

Molemmat "sip:"- ja "sips:"-skeemat noudattavat RFC 2396 –määrittelyn sääntöjä. Ne muodostuvat samankaltaisesti kuin ”mailto” Internet-resurssin tunniste. Muodollinen syntaksi SIP Internet-resurssin tunnisteelle on

"sip:user:password@host;port;uri-parameters?headers". SIPS Internet-resurssin tunnisteen syntaksi on sama, mutta se alkaa "sip":n sijaan "sips".

Käyttäjän identifioi toimialueella "user"-kenttä. Mikäli palvelin jolle pyyntö lähetetään, osaa käsitellä puhelinnumeroita, kuten ”Internet telephone gateway”, puhelinnumero voi sijaita "user"-kentässä. Käyttäjän salasana voidaan välittää "password"-kentässä. Tätä ei kuitenkaan suositella käytettäväksi, koska salasana lähetetään tällöin selko-kielisenä (clear text). Näin voitaisiin lähettää esimerkiksi puhelimen pin-numero selkokielisenä. Toimialue, jolla SIP-resurssi sijaitsee on "host"-kentässä. Kenttä koostuu joko verkko-tunnuksesta tai numeerisesta IPv4- tai IPv6-osoitteesta. Portti, johon pyyntö lähetetään ilmaistaan "port"-kentässä. Ylimääräisiä parametrejä voidaana välittää "uri-parameters"-kentässä. Parametrit erotellaan puolipistein ja annetaan muodossa parametrin_nimi=parametrin_arvo. Otsakenttiä voidaan antaa "headers"-kentässä. Otsakekentät erotetaan ”&”-merkillä ja annetaan muodossa otsakekentän_nimi=kentän_arvo. Erityisellä "body"-otsakekentällä voidaan määrittää SIP-viestin sisältö. Esimerkkejä SIP- ja SIPS-osoitteista on esitetty taulukossa 1.

Taulukko 1. SIP- ja SIPS-osoite-esimerkkejä

sip:petri@uta.fi
sip:petri:secretworld@uta.fi;transport=tcp
sips:petri@uta.fi?subject=project%20x&priority=urgent
sip:+358-50-123456:1234@gateway.com;user=phone
sips:1212@gateway.com
sip:petri@192.0.2.4
sip:uta.fi;method=REGISTER?to=petri%40uta.fi

3.6. SIP-viestit

SIP-viestit jaetaan pyyntöihin ja vasteisiin. Viestit koostuvat neljästä osasta: aloitusrivistä, yhdestä tai useammasta otsikkorivistä, tyhjästä rivistä ja viestin rungosta. Viestin jokaisen rivin tulee päättyä rivinvaihtoon. Aloitusrivi sisältää joko pyyntörivin tai vasterivin. Viestillä ei tarvitse olla runkoa, mutta tyhjä rivi otsikkorivin tai otsikkorivien jälkeen on pakollinen. Viestin rakennetta on havainnollistettu koodiesimerkissä 3, jossa on esitetty yhteyden luonnissa käytettä INVITE-viesti. Viestien muoto noudattaa Internetin standardoitua viestiformaattia [RFC 2822].

INVITE sip:bob@biloxi.com SIP/2.0	Aloitusrivi
Via: SIP/2.0/UDP pc33.atlanta.com;branch=z9hG4bK776asdhds Max-Forwards: 70 To: Bob <sip:bob@biloxi.com> From: Alice <sip:alice@atlanta.com>;tag=1928301774 Call-ID: a84b4c76e66710@pc33.atlanta.com CSeq: 314159 INVITE Contact: <sip:alice@pc33.atlanta.com> Content-Type: application/sdp Content-Length: 142	Otsikkorivit
	Tyhjä rivi
(SDP viesti, jota ei tässä näytetä)	Viestin runko

Koodiesimerkki 3. SIP-viestin rakenne

SIP-tapahtumaksi kutsutaan tiettyyn viestiin liittyvää viestinvaihtoa asiakkaan ja palvelimen välillä. Tapahtuma sisältää aina pyynnön ja lopullisen vasteen. Tapahtuma voi myös sisältää yhden tai useampia tilapäisiä vasteita.

SIP-viestin ollessa pyyntö, se sisältää tilatiedon jossa halutaan jokin toiminne toteutettavaksi. Sanoman kulkusuunta on asiakkaalta palvelimelle. SIP-pyyntö tunnistetaan pyyntörivistä (koodiesimerkki 4), joka sisältää metodin nimen, pyynnön vastaanottajan URI-osoitteen ja protokollan version. SIP-määrittely määrittelee kuusi eri SIP-metodia: INVITE, ACK, OPTIONS, BYE, CANCEL ja REGISTER (taulukko 2). Lisäksi SIP:in lisäosat voivat määritellä lisää metodeja [RFC 3261].

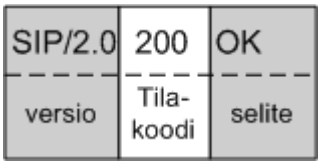
Koodiesimerkki 4. SIP-pyyntörivin osat

Taulukko 2. SIP-metodit

SIP-metodi	Selite
INVITE	INVITE-pyyntö kutsuu käyttäjän istuntoon, viestin runko sisältää istunnon kuvauksen.
ACK	ACK-viesti lähetetään lopullisen viestin vastaanottamisen varmistamiseksi. Asiakas lähettää INVITE-pyynnön ja saatuaan vasteen pyyntöön hän lähettää ACK-viestin. Näin saavutetaan kolmiosainen kättely: INVITE – lopullinen vaste – ACK.
CANCEL	CANCEL-pyyntö peruuttaa toteutumattoman tapahtuman. Mikäli INVITE-pyyntö on lähetetty, mutta lopullista vastausta ei ole vielä saatu, INVITE-pyynnön käsittely lopetetaan.
BYE	BYE-pyynnöllä käyttäjä poistuu istunnosta.
REGISTER	REGISTER-pyynnöllä käyttäjä ilmoittaa palvelimella nykyisen sijaintinsa.

OPTIONS	OPTIONS-pyynnöllä käyttäjä kysyy palvelimelta sen ominaisuuksista, kuten mitä metodeja ja mitä istunnon kuvauskieliä se tuntee.

Jokaisen pyynnön kohdalla palvelin luo vasteen. Vasteen kulkusuunta on palvelimelta asiakkaalle, ja se sisältää tilatiedon pyynnön edellyttämistä toiminnoista. Jokaisella vasteella on sen tilan ilmaiseva tilakoodi. SIP-vaste tunnistetaan vasteen tilarivistä (koodiesimerkki 5). Vasteen tilarivi sisältää protokollan version, tilakoodin ja tekstimuotoisen seliteosan [RFC 3261].



Koodiesimerkki 5. SIP-vasteen tilarivin osat

Tilakoodit ovat kokonaislukuja väliltä [100…699]. Ne on luokiteltu (taulukko 3) käyttötarkoituksensa mukaisesti. Vasteet väliltä [100…199] ovat tilapäisiä, vastaavasti vasteet väliltä [200…699] ovat lopullisia. Vain lopulliset vasteet voivat lopettaa SIP-tapahtuman, tilapäiset vasteet jättävät tapahtuman avoinaiseksi odottamaan myöhemmin lähetettävää lopullista vastetta [RFC 3261].

Taulukko 3. SIP-vasteiden tilakoodit

Koodialue	**Vasteen tyyppi**	**Merkitys**
100-199	Informational	Pyyntö vastaanotettu, jatketaan pyynnön prosessointia.
200-299	Success	Pyyntö on suoritetu onnistuneesti.
300-399	Redirection	Pyynnön toteuttamiseksi tarvitaan lisätoimenpiteitä.
400-499	Client error	Pyyntö on virheellinen ja sitä ei voida suorittaa.
500-599	Server error	Palvelin epäonnistui oikeanmuotoisen pyynnön suorittamisessa.
600-699	Global failure	Pyyntöä ei voida suorittaa millään palvelimella.

Jokaiseen SIP-viestiin kuuluva aloitusrivi kertoo mistä viestissä on kyse. Osa viestiin kuuluvista otsikkokentistä on useammin käytettyjä kuin toiset, itse asiassa ne sisältyvät melkein kaikkiin viesteihin. Taulukossa 4 on esitetty yleisimmät otsakekentät ja niiden merkitykset.

Taulukko 4. SIP-viestien yleisimmät otsikkokentät

Kentän nimi	**Kentän merkitys**
Via	Kertoo minkä palvelimen kautta yhteys muodostetaan. Välipalvelimet lisäävät näitä kenttiä. Vastaanottaja näkee Via-kentistä mitä reittiä viesti on kulkenut.
From	Viestin lähettäjän osoite.
To	Viestin vastaanottajan osoite.
Call-ID	Sisältää paikallisesti yksilöidyn ja globaalisti yksilöidyn viestinumeron.

CSeq	Sisältää sanoman yksilöintitunnisteen istunnon aikana.
Contact	Sisältää yhteystiedon kyseiselle yhteydelle.
Content-Length	Sanoman pituus otsikko-kenttien jälkeen.

SIP on tekstipohjainen eikä sitä ole suunniteltu erityisen tehokkaaksi protokollaksi. Alunperin se suunniteltiin luomaan istuntoja, joissa välitettäisiin multimedia-dataa, joten tehokkuus ei ollut tärkein kriteeri. Tilanne kuitenkin muuttui, kun SIP:iä alettiin käyttämään signalointiprotokollana langattomissa verkoissa. Viestinnän tehokkuuteen pitää tällöin kiinnittää erityistä huomiota. Tätä varten kehitettiin SigComp, joka mahdollistaa viestien pakkaamisen tehokkaasti ja dataa menettämättä [RFC 3320].

3.7. Rekisteröinti

Rekisteröinti on tapahtuma, jonka avulla käyttäjä ilmoittaa sijaintinsa rekisteröintipalvelimelle. Välipalvelimet tarvitsevat tätä tietoa, jotta ne voivat lähettää käyttäjälle tulevat SIP-viestit käyttäjän päätelaitteeseen. Käyttäjän päätelaite lähettää näitä rekisteröintipyyntöjä tai -ilmoituksia tietyin väliajoin, jotta rekisteröintipalvelimella oleva tieto pysyy ajan tasalla. REGISTER-viesti liittää käyttäjän SIP-osoitteen (esim: sip:petri.lintula@uta.fi) käyttäjän päätelaitteen ip-osoitteeseen. Rekisteröintipalvelin hyväksyy vain REGISTER-viestejä ja hylkää muut SIP-viestit. Rekisteröintipalvelin tallentaa saamansa tiedon, jota kutsutaan myös nimellä sidonta (binding), paikkatietopalveluun. Usein rekisteröintipalvelin on fyysisesti sama palvelin kuin välipalvelin. On kuitenkin tärkeää erotella nämä kaksi eri "palveluiksi" [RFC 3261].

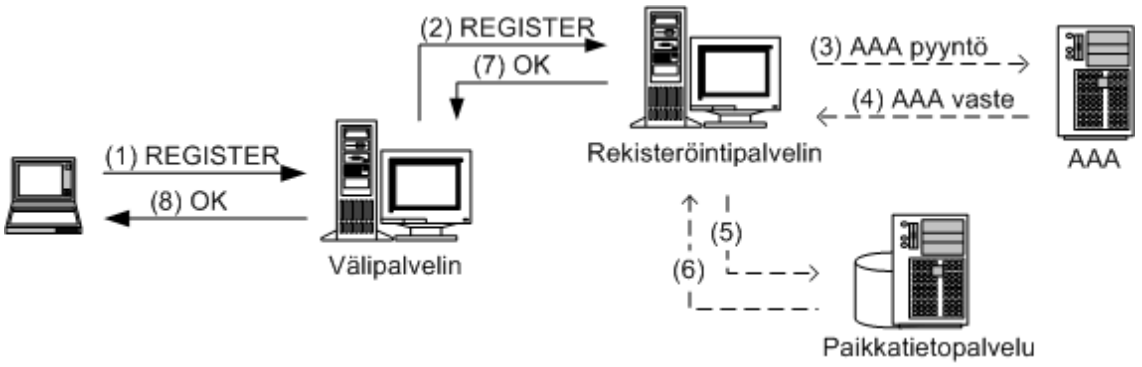
Paikkatietopalvelu on pelkkä abstrakti käsite. SIP ei vaadi tiettyä toteutustapaa sen toteuttamiseksi. Ainoat vaatimukset ovat, että jonkun toimialueen rekisteröintipalvelu pystyy lukemaan ja kirjoittamaan tietoja paikkatietopalveluun sekä välipalvelimen ja uudelleenohjauspalvelimen pitää pystyä lukemaan dataa paikkatietopalvelusta.

Rekisteröinnin tarkoituksena on pelkästään löytää käyttäjän päätelaite, eikä sillä ei ole mitään roolia valtuuttaa viestien välitystä. Valtuutus (authorization) ja todennus (authentication) käsitellään SIP:issä joko pyyntö kerrallaan pyyntö-vaste -mekanismilla (challenge-response) tai käyttämällä alemman tason mekanismeja.

3.7.1. Sidontojen muokkaaminen

Käyttäjä luo uuden sidonnan lähettämällä rekisteröintipalvelimelle REGISTER-viestin. Palvelin tallentaa viestissä olevat tiedot paikkatietopalveluun ja lähettää käyttäjälle kuittauksena SIP OK –viestin,

jonka seliteosa sisältää kaikki käyttäjän voimassa olevat sidonnat. Rekisteröintiprosessin kulku on esitetty kuvassa 8. Kuvassa rekisteröintipalvelin lisäksi tarkistaa onko käyttäjällä oikeus muokata sidontoja.



Kuva 8. Rekisteröintiprosessin kulku

Käyttäjä voi rekisteröidä useita päätelaitteita rekisteröintipalvelimelle. Näin välipalvelin voi tavoittaa käyttäjän useasta eri paikasta. On myös mahdollista että useampi käyttäjä rekisteröityy käyttämään samaa päätelaitetta.

Käyttäjäagentilla on kolme tapaa päätellä minne se lähettää rekisteröintejä: konfiguroinnin kautta, käyttämällä tunnettua osoitetta ja joukkolevityksellä (multicast). Käyttäjäagentti voidaan konfiguroida käyttämään aina tiettyjä rekisteröintipalvelimia. Mikäli konfiguroituja rekisteröintipalvelimia ei löydy, päätelaite lähettää pyynnön toimialueen ”isäntäkoneelle” käyttäen dokumentissa RFC 3263 määriteltyjä SIP-palvelinten etsintämekanismeja. Esimerkiksi käyttäjän sip:petri.lintula@uta.fi päätelaite lähettää pyynnön osoitteeseen sip:uta.fi. Lopuksi päätelaite voidaan konfiguroida käyttämään joukkolevitystä (multicast), jolloin rekisteröinti osoitetaan tunnetuille SIP-palvelimille joukkolevitys-osoitteeseen "sip.mcast.net" (224.0.1.75).

Sidontojen muokkaaminen tapahtuu lähettämällä REGISTER-viesti rekisteröintipalvelimelle, joka päivittää sidontoja saamansa REGISTER-viestin pohjalta. REGISTER-viesti sisältää yhden tai useamman yhteystiedon, joka kertoo minne käyttäjä haluaa hänelle tulevat SIP-viestit ohjata. Käyttäjän osoite on näkyvissä REGISTER-viestin “to”-kentässä. “Contact”-kenttä muodostuu SIP tai SIPS Internet-resurssin tunnisteista, jotka yksilöivät pyynnön loppukohteen (esimerkiksi sip:pl69881@mobile123a.uta.fi). SIP-päätelaite voi rekisteröidä myös puhelinnumeroita käyttäen RFC 2806 –määrittelyn sääntöjä, tai sähköpostiosoitteita käyttäen RFC 2368 –määrittelyn sääntöjä ”Contact”-kentän arvoina. Esimerkiksi ”petri” osoitteesta "sip:petri@uta.fi", rekisteröityy toimialueen uta.fi rekisteröintipalvelimelle. Hänen sidontojaan käyttäisivät

tällöin uta.fi –toimialueen välipalvelimet, jotta ne osaisivat välittää viestejä Petrin päätelaitteeseen. Kun käyttäjä on luonut sidonnan rekisteröinti-palvelimelle, hän voi lähettää uusia REGISTER-viestejä, joissa on tietoja uusista sidonnoista tai muokata olemassa olevia sidontoja.

Käyttäjäagentti voi ehdottaa REGISTER-viestille umpeutumisaikaa (expiration interval), jonka jälkeen rekisteröinti ei enää ole voimassa. Rekisteröintipalvelin kuitenkin valitsee todellisen umpeutumisajan asetustensa perusteella. Rekisteröinnin umpeutumisen voi välttää lähettämällä uuden REGISTER-viestin rekisteröintipalvelimelle. Jokainen käyttäjä on itse vastuussa sidontojensa päivittämisestä. Käyttäjä saa REGISTER-viestiin vastaukseksi SIP OK –viestin, joka sisältää kaikki käyttäjän voimassa olevat sidonnat. Käyttäjä voi pyytää välitöntä sidonnan poistamista määrittelemällä umpeutusmisajaksi nollan (0) REGISTER-viestissä.

Jos REGISTER-viesti sisältää enemmän kuin yhden ”Contact”-osoitteen, niin ne voidaan priorisoida "q"-parametrin avulla. Näin käyttäjä voi asettaa päätelaitteensa hänelle mieluisaan järjestykseen.

3.7.2. REGISTER-viesti

REGISTER-viestit lisäävät, poistavat ja kyselevät sidontoja. Yksi REGISTER-viesti voi lisätä yhden tai useamman uuden sidonnan palveluun samalla kertaa. Valtuutettu kolmas osapuoli voi myös luoda rekisteröinnin. Käyttäjä voi myös poistaa aikaisemman rekisteröinnin tai kysellä mitkä sidonnat ovat voimassa.

REGISTER-viestin tulee sisältää seuraavat kentät: Request-URI, To, From, Call-ID, Cseq sekä Contact. ”Request-URI”-kenttä määrittää toimialueen jolle rekisteröinti on tarkoitettu (esim. sip:uta.fi). ”To”-kenttä määrittää käyttäjä-nimen, jota rekisteröinti koskee. ”From”-kenttä sisältää tiedon siitä kuka rekisteröinnin tekijä on. Tämä on sama kuin ”to”-kenttä, ellei rekisteröintiä tee kolmas osapuoli. ”Call-ID”-kenttä, identifioi päätelaitteen. Sen tulee olla sama kaikissa REGISTER-viesteissä yhdeltä päätelaitteelta. ”Cseq”-kenttä varmistaa REGISTER-viestien asettamisen oikeaan järjestykseen. Päätelaite kasvattaa kentän arvoa yhdellä jokaista pyyntöä kohden. “Contact”-kenttä voi olla sisältämättä sidontoja tai sisältää useita sidontoja.

Sidonnoille voidaan myös ehdottaa umpeutumisaikaa. Tämä tapahtuu ”Expires”-otsakekentällä tai ”expires”-parametrillä ”Contact”-kentässä. ”Contact”-kentän parametri mahdollistaa umpetumisajan määrittämisen jokaiselle sidonnalle erikseen, kun taas ”Expires”-otsakekentällä määritetään umpeutusmisaika kaikille ”Contact”-kentän osoitteille, joilla ei ole "expires"-parametriä. Näin on mahdollista määrittää eri umpeutusmisajat eri päätelaitteille. Koodiesimerkissä 6 on esitetty esimerkki REGISTER-viestistä.

```
REGISTER sip:registrar.foo.com SIP / 2.0
Via: SIP / 2.0 / UDP kayttajapc.foo.com:5060;branch=z9hG4bKnashds7
Max-Forwards: 70
Request-URI: sip:foo.com
To: Kayttaja <sip:kayttaja@foo.com>
From: Kayttaja <sip:kayttaja@foo.com>;tag=456248
Call-ID: 843817637684230@998sdasdh09
CSeq: 1826 REGISTER
Contact: <sip:kayttaja@192.0.2.4>
Expires: 7200
Content-Length: 0
```

Koodiesimerkki 6. REGISTER-viesti

4. Suoraviestintä

Suoraviestintä tarkoittaa käyttäjien välisten viestien lähetystä lähes reaaliaikaisesti. Viestit ovat yleensä lyhyitä tekstiviestejä ja viestintä on usein nopeatempoista. Viestejä ei yleensä tallenneta, mutta se on mahdollista. Suoraviestintä eroaa sähköpostista käyttötavaltaan, useiden lyhyiden viestien lähettäminen edestakaisin muistuttaa paljolti keskustelua.

Suoraviestintä itsessään on ollut olemassa jo jonkin aikaa. Aikaisempia suoraviestintä toteutuksia ovat muunmuassa zephyr ja IRC. Viime aikoina suoraviestinnän yhteyteen on liitetty läsnäolotieto ja kaverilistat (buddy list). Käyttäjä saa ilmoituksen kaverilistalla olevan henkilön kirjautuessa järjestelmään. Tähän asti suoraviestinnän toteutustavat ovat olleet sovelluskohtaisia, joten eri suoraviestintäsovellukset eivät ole voineet toimia keskenään. SIP yrittää tuoda tähän muutosta. IETF on määritellyt yleiset vaatimukset sekä mallin läsnäolo ja suoraviesti protokollille, jotka on kuvattu määrittelyissä RFC 2778 ja RFC 2779.

Suoraviestintä SIP:issä tapahtuu käyttäen SIMPLE:ä (SIP for Instant Messaging and Presence Leveraging Extensions). Sitä kehittää IETF:n SIMPLE-työryhmä, joka keskittyy kehittäämään SIP:in ympärille standardia suoraviestinnälle ja läsnäololle. Sen julkaisemat SIP:in laajennukset ovat mahdollistaneet suoraviestien lähetyksen (RFC 3428) sekä läsnäolotapahtumien välittämisen (RFC 3265) SIP:illä.

SIP mahdollistaa suoraviestinnän kahdella eri tavalla: pager-moodilla (pager mode) sekä istunto-moodilla (session mode). Pager-moodissa [RFC 3428] jokainen viesti on itsenäinen ja sisältää kaiken viestimisen tarvitseman informaation. Tämä lähetystapa on hyvä käytettäessä lyhyitä viestijaksoja, lyhyitä viestejä tai lähetettäessä viesti useille käyttäjille. Istunto-moodi [Campbell et al.] vastaavasti on joukko viestejä; SIP:iä käytetään tällöin luomaan yhteys kahden eri päätelaitteen välille. Toinen protokolla välittää viestit. Parhaiten tälläiseksi protokollaksi soveltuu TCP/SCTP. Tämä lähetystapa on hyvä silloin kun lähetetään useita viestejä tai suuria viestejä. Moodia voidaan laajentaa juttutorin (chat room) tapaiseksi sessioksi, jossa viestit välitetään suurelle joukolle käyttäjiä tai käyttäjäjoukolle, jonka sisällä vaihtuvuus on suuri.

Lyhyille viesteille pager-moodi on tehokkaampi, erityisesti langattomassa ympäristössä. Se mahdollistaa matalamman saantiviiveen (latency) käyttäjälle kuin erillisen viestintäistunnon luominen. Mikäli käyttäjä aikoo lähettää useita viestejä tai suuria viestejä, istunto-moodi on tehokkaampi. Viesti voi sisältää

viitteen (esimerkiksi URL-osoitteen), joka osoittaa suureen viestiin sen sijaan että se sisältäisi itse viestin.

Suoraviestintää tullaan käyttämään myös 3G-verkoissa, joissa sen välittämisestä vastaa IP Multimedia Subsystem eli IMS. IMS on osa 3G-verkkoa ja se on täysin IP-pohjainen.

3G-verkoissa suoraviestinnän tulee olla mahdollista kahdella eri tavalla: välitön viestintä (immediate messaging) ja istuntopohjainenviestintä (session based messaging). Välittömässä viestinnässä lähettäjä odottaa viestin menevän perille ja luettavan lähes reaaliaikaisesti. Lähettäjä on tällöin yleensä tietoinen vastaanottajan läsnäolotilasta läsnäolotietopalvelun kautta. Istuntopohjaisessa viestinnässä lähettäjä ja vastaanottaja perustavat istunnon viestintää varten. Molempien viestintätyyppien vaatimukset ovat hyvin samanlaisia kuin SIP:in tarjoamat mahdollisuudet. Välitöntä viestintää vastaa pager-moodin viestintä ja istuntopohjaistaviestintää istuntoviestintä [Niemi].

3G-verkko asettaa viestinnälle omat erikoisvaatimuksensa. Viestit välitetään langattomasti, joten viestinnän on oltava tehokasta ja turhaa viestintää vältettävä. Tässä SIP:in tekstipohjaisuus kääntyy sitä vastaan, koska se ei ole erityisen tehokas. Lisäksi päätelaitteiden resurssit voivat olla hyvinkin pienet: muistin määrä ja prosessointitehon voivat olla vaatimattomat, suuri virrankulutus ei ole mahdollista ja näyttö asettaa omat rajoituksensa värien ja resoluution muodossa [Niemi].

4.1. Päätelaitteet

Käyttäjä ei välttämättä tiedä viestin vastaanottajan olinpaikkaa eikä hänen päätelaitettaan. Erot päätelaitteiden ominaisuuksissa saattavat johtaa siihen, ettei vastaanottaja pystykään lukemaan viestiä tai jotain sen osaa. OPTIONS-viesti mahdollistaa vastaanottajan päätelaitteen ominaisuuksien kyselyn, mutta käyttäjää ei voi vaatia kysymään jokaisen vastaanottajan päätelaitteen ominaisuuksia.

Kaikkien päätelaitteiden tulee pystyä käsittelemään tekstimuotoisia viestejä. Viestien sisältöjen ollessa suuria ja muodostuessa eri mediatyypeistä, kuten kuvista ja videoleikkeistä, ei voida olettaa kaikkien päätelaitteiden osaavan käsitellä niitä. Erityisen haastavaa tämä on mobiilipäätelaitteiden kohdalla, jolloin mukaan tulee muita rajoituksia kuten näytön koko ja resoluutio, muistin määrä sekä tuetut viestiformaatit. Esimerkiksi viesti voi olla liian suuri mahtuakseen päätelaitteen muistiin.

Päätelaiteriippumattomuuden saavuttamiseksi on ehdotettu että viestit muutetaan vastaanottajan päätelaitteelle sopivaksi. Tämä muunnos voi tapahtua lähettäjän päätelaitteessa tai SIP-palvelimella, jolloin lähettäjän ei tarvitse huolehtia viestin vastaanottajien päätelaitteiden ominaisuuksista.

Vastaanottaja taasen saa viestin, joka on muokattu hänen päätelaitteelleen, eikä kaikille vastaanottajille sopivaan formaattiin. Kaikki päätelaitteet eivät välttämättä tue SIP:iä, eivätkä kaikki viestit ole SIP-viestejä. Tämän ongelman ratkaisemiseksi IETF on kehittämässä yhteistä formaattia suoraviestinnälle nimeltä CPIM (Common Profile for Instant Messaging) [CPIM].

4.2. Pager-moodi

SIP:iin on tehty laajennos, SIMPLE, joka mahdollistaa suoraviestien lähetyksen. Se perii automaattisesti SIP:in ominaisuudet, koska kyseessä on laajennos. Suoraviestin lähetykseen pager-moodissa käytetään MESSAGE-viestiä. Itse suoraviesti sijoitetaan MESSAGE-viestin runkoon, josta vastaanottaja sen tulkitsee. Viestin sisältö voi olla mitä tahansa MIME-koodattua sisältöä, mutta yleensä sisältö on tekstiä [RFC 3428].

Jokaista MESSAGE-viestiä käsitellään yksittäisenä viestinä. Peräkkäisistä viesteistä ei teknisesti muodostu keskustelua, vaikka näin käyttäjä ehkä luuleekin. Pager-moodi muistuttaa hyvin paljon matkapuhelimilla lähetettyjä sms-viestejä.

Toisin kuin istunto-moodissa, pager-moodin viestillä voi olla vain yksi vastaanottaja. Pager-moodi ei mahdollista juttutorin tapaisia ratkaisuja. Sama viesti voidaan lähettää usealle eri käyttäjälle, mutta silloin viesti lähetetään jokaiselle käyttäjälle erikseen.

Kuvassa 9 on esitetty MESSAGE-viestin kulku. KäyttäjäA lähettää MESSAGE-viestin KäyttäjäB:lle. Viesti voi mennä suoraan toiselle käyttäjälle, mutta yleensä se kulkee yhden tai useamman SIP-palvelimen kautta. MESSAGE-viestin saatuaan vastaanottaja lähettää OK-viestin lähettäjälle. Näin saadaan varmistus viestin perillemenosta. Viestin lukemisesta ei kuitenkaan voida varmistua. Kuvassa ei ole esitetty vaihetta, jossa välipalvelin kyselee paikkatietopalvelulta, mistä KäyttäjäB löytyy.

Kuva 9. MESSAGE-viestin välitys käyttäjien välillä

Kuvassa 9 lähetettävän MESSAGE-viestin muoto on esitetty koodiesimerkissä 7. Vastaavasti OK-viestin muoto on esitetty koodiesimerkissä 8. Viestin vastaanottajan osoite on määritelty suoraviestintä Internet-resurssin nimellä (IM URI) muodossa im:user@domain. Nämä

osoitteet ovat abstrakteja ja ne ovat yleensä muunnettu SIP-osoitteiksi CPIM-määritelmän mukaisesti [CPIM].

```
MESSAGE im:KayttajaB@domain.com SIP/2.0
Via: SIP/2.0/TCP 1.2.3.4:5060
From: sip:KayttajaA@domain.com
To: sip:KayttajaB@domain.com
Contact: sip:smsClient@1.2.3.4
Call-ID: smsClient481948@1.2.3.4
CSeq: 1 MESSAGE
Content-Type: text/plain
Content-Length: 11
<tyhjä rivi>
Hello world
```

Koodiesimerkki 7. (1) MESSAGE-viesti

```
SIP/2.0 200 OK
Via: SIP/2.0/TCP 1.2.3.4:5060
From: sip:KayttajaB@domain.com
To: sip:KayttajaA@domain.com
Call-ID: smsClient481948@1.2.3.4
CSeq: 1 MESSAGE
Content-Length: 0
```

Koodiesimerkki 8. (3) SIP OK –viesti

MESSAGE-viesti poikkeaa muista SIP-viesteistä, koska siinä välitetään myös tietoa. Muita SIP-viestejä käytetään pääsääntöisesti pelkästään signalointiin. Viestit lähetetään normaalisti selkokielisinä, joten tietoturva korostuu. Kuka tahansa pystyy lukemaan viestin kaappaamalla verkkoliikennettä. Paras keino suojata viesti on salata se käyttäen S/MIME-mekanismia.

4.3. Istunto-moodi

Suoraviestinnälle on olemassa istunto-moodi, jossa yksittäinen viesti kuuluu aina johonkin istuntoon. Useita viestejä lähetettäessä sen avulla voidaan huomattavasti vähentää liikennettä välipalvelinten välillä. Jokainen pager-moodin viesti vaatii kokonaisen SIP-tapahtuman, eli pyynnön ja vasteen. Istunto-moodin luonti vaatii 5 SIP-viestiä eli kolmen tai useamman pager-moodi viestin lähettäminen vaatii enemmän viestejä. Istunnon luonnin jälkeen viestit välittyvät suoraan vastaanottajalle eivätkä enää kulje välipalvelinten kautta. Lisäksi pager-moodissa olevia otsaketietoja ei yleensä tarvita istunto-moodissa sen jälkeen kun istunto on luotu. Istuntomoodissa ei myöskään ole

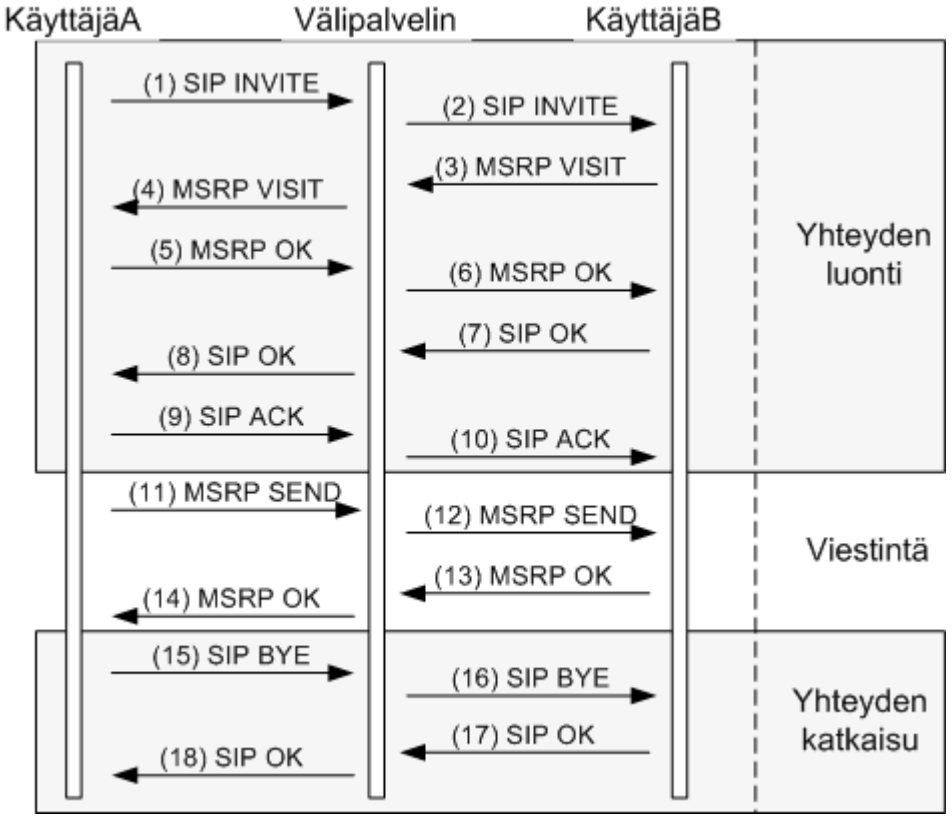
mitään rajoituksia viestin koolle. Se ei vaadi kuittausta viestin perille menosta ennen kuin se voi lähettää uuden viestin [Cambpell et al.].

Viestien salaaminen on tehokkaampaa istunto-moodissa. Pager-moodissa viestin sisältö pitää salata S/MIME-toiminnoilla, joka vaatii julkisen avaimen lähettämisen kaikkien viestien yhteydessä. Istunto-moodissa voidaan luoda istuntokohtainen avain salaamaan jokainen viesti.

Message Session Relay Protocol (MSRP) on tekstipohjainen viestien välitykseen tarkoitettu protokolla. Se tarjoaa mekanismin välittää viestejä istunto-moodissa. MSRP-viestit muodostuvat samaan tapaan kuin SIP-viestit, ne ovat joko pyyntöjä tai vasteita. MSRP määrittelee viestit SEND ja VISIT. SEND-viestin avulla välitetään itse viestejä päätelaitteiden välillä, VISIT-viestillä luodaan istunto toiseen käyttäjään. MSRP-tapahtuma sisältää SIP-tapahtumasta poiketen vain yhden pyynnön ja yhden vasteen.

Kuvassa 10 on esitetty viestin lähetys istunto-moodissa. Kuvassa on myös eritelty yhteyden luonti-, viestintä- sekä yhteyden katkaisuosat. Luonnollisesti viestintäosassa on mahdollista lähettää useita viestejä, esimerkissä lähetetään kuitenkin vain yksi viesti.

KäyttäjäA lähettää SIP INVITE –viestin (koodiesimerkki 9) KäyttäjäB:lle luodakseen istunnon. KäyttäjäB vastaa MSRP VISIT –viestillä luodakseen TCP yhteyden KäyttäjäA:n INVITE-viestissä määrittämään osoitteeseen ”kayttajaA.domain1.com :7777/pc01” (koodiesimerkki 10), johon hän saa vastaukseksi MSRP OK- viestin täydentäen näin MSRP-tapahtuman. Tämän jälkeen hän lähettää KäyttäjäA:lle SIP OK –viestin, että istunto on luotu. Itse viestien välitys tapahtuu MSRP SEND –viesteillä (koodiesimerkki 11), jonka toinen käyttäjä aina kuittaa vastaanotetuksi MSRP OK –viestillä (koodiesimerkki 12). Istunnon lopetus tapahtuu SIP BYE –viestillä.



Kuva 10. Suoraviestintä istunto-moodissa

```
INVITE sip:kayttajaB@domain2.com
v=0
o=kayttajaA 2890844557 2890844559 IN IP4 host.anywhere.com
s=
c=IN IP4 fillername
t=0 0
m=message 9999 msrp/tcp *
a=accept-types:text/plain
a=direction:both 0
a=session:msrp://kayttajaA.domain1.com:7777/pc01
```

Koodiesimerkki 9. (1) SIP INVITE-viesti

```
MSRP xx VISIT
S-URL:msrp://kayttajaA.domain1.com:7777/pc01
Tr-ID: sie09s
```

Koodiesimerkki 10. (3) MSRP VISIT-viesti

```
MSRP xx SEND
TR-ID: 123
Content-Type: "text / plain"
Hello world!
```

Koodiesimerkki 11. (11) MPRS SEND-viesti

```
MSRP xx 200 OK
TR-ID: 123
```

Koodiesimerkki 12. (13) MPRS OK-viesti

4.4. Suoraviestintä yrityksissä

Useat yritykset ovat huomanneet suoraviestinnän tehostavan niiden toimintaa. Suoraviestintää ei enää pidetä nuorten keskusteluna vaan yhtenä lisäkeinona kommunikoida työtoverin tai asiakkaan kanssa.

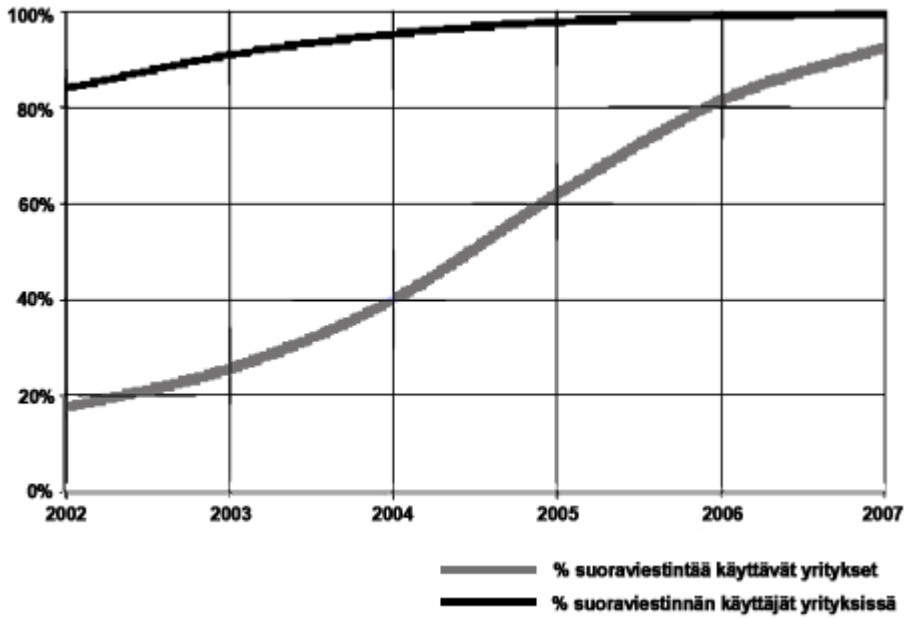
Osterman Research –tutkimusyrityksen mukaan vuonna 2002 42 % yrityksistä käytti suoraviestintää osana virallisia työskentelytapojaan. Luvun odotettiin kasvavan 65 – 70 %:iin vuonna 2003. Taulukossa 5 on esitetty tutkimusyrityksen toteuttama sama tutkimus vuosina 2001 - 2003. Tutkimuksiin osallistui hieman alle 200 yritystä. Syyskuussa 2002 tehdyssä tutkimuksessa todettiin myös, että epävirallisesti 84 %:ssa yrityksistä käytetään suora-viestintää, ja tämän luvun odotettiin kasvavan 93 %:iin seuraavan 12 kuukauden aikana. Helmikuussa 2003 tehdyssä tutkimuksessa lähes kaikkien yritysten odotettiin käyttävän suoraviestintää vuonna 2005 (kuva 12) [Osterman Research].

Taulukko 5. Suoraviestinnän käyttö yrityksissä

	Heinäkuu 2002	**Maaliskuu 2002**	**Syyskuu 2002**	**Syyskuu 2003**	**Maaliskuu 2004**
Käyttää suoraviestintää	21	29	42	44	44
Ei käytä, mutta aikoo käyttää	11	11	8	7	13
Ei käytä, mutta ehkä aloittaa käytön	40	31	28	30	24
Ei aio aloittaa käyttöä	28	29	22	19	19

Tutkimuksissa kysyttiin myös tietohallinnon mielipidettä suoraviestinnästä. Ensimmäisessä vuonna 2001 tehdyssä tutkimuksessa 22 % yritysten tietohallin-noista kannatti suoraviestintää osana yrityksen kommunikointitapoja. Vuonna 2003 tehdyssä tutkimuksessa vastaava luku oli 39 prosenttia. Tämä johtuu suuresti siitä, että ensimmäisessä tutkimuksessa suurimmalla osalla yrityksistä ei ollut mielipidettä suoraviestinnästä. Myös

suoraviestinnän vastustajien kannatus on lisääntynyt samana aikana 17 %:sta 23 %:iin.



Kuva 11. Suoraviestinnän markkoiden kasvu

Tällä hetkellä useat yritykset ovat siirtymässä käyttämään omia suoraviestintäfoorumejaan. Vastaavasti kuin yrityksellä on oma sähköposti-palvelin, niin heillä on myös oma suoraviestintäpalvelin. Usein suoraviestintä toteutetaan yrityksen ekstranet-palveluun keskusteluhuoneen tapaiseksi palveluksi. Tämä mahdollistaa sen, että viestit säilyvät ja käyttäjät voivat lukea myös aikaisemmin lähetettyjä viestejä. Hallinnoidessaan suoraviestintä-palveluaan yritykset voivat päättää kenelle he antavat oikeudet foorumeihinsa. Lisäksi valtuuttamaton foorumien käyttö tulee mahdottomaksi, koska yritys tuntee kaikki asiakkaan käyttäjät ja he vastaavasti tuntevat yrityksen käyttäjät. Julkisissa palveluissa (esimerkiksi AOL Instant Messengerissä) ulkopuolisten on mahdollista osallistua keskusteluun keskusteluhuoneessa. Lisäksi julkisissa palveluissa käyttäjien kutsumanimet ovat usein vähemmän virallisia, eikä välttämättä oikeita yhteystietoja edes haluta käyttää.

5. Läsnäolo

Läsnäolo kertoo käyttäjän mahdollisuuden ja halukkuuden kommunikoida käyttäen useita eri päätelaitteita, kommunikointitapoja ja mediatyyppejä. Läsnäoloon vaikuttaa moni asia; esimerkiksi onko henkilön matkapuhelin päällä, puhuuko hän puhelua, ilmoittaako kalenteriohjelma henkilön olevan kokouksessa, onko suoraviestintäohjelma käynnissä ja ilmoittaako se henkilön olevan lounaalla. Lisäksi läsnäoloon kuuluu henkilön halu olla tavoitettavissa (availability). Läsnäolotieto luullaan usein kuuluvan osaksi suoraviestintäsovellusta. Näin ei kuitenkaan ole, vaan se on siitä erillään.

Liittämällä läsnäolotieto suoraviestisovellukseen parannetaan huomattavasti sovellusten hyödyllisyyttä. Esimerkiksi SMS-viestit ovat aika- ja paikkariippumattomia. Lähettäjä ei kuitenkaan voi olla varma, saavuttiko viesti vastaanottajan vai vain hänen puhelimensa. Läsnäolotiedolla pyritään poistamaan tämä ongelma. Käyttäjä voi itse asettaa oman läsnäolotilansa, esimerkiksi ”lounaalla”, ”työpaikalla”, ”kokouksessa” tai ”tavoittamattomissa”. Lähettäessään tiedon palvelimelle muut käyttäjät tietävät, onko käyttäjä tavoitettavissa. Kaverilistoihin voi kerätä useita ystäviä, jolloin sovellus tarkkailee listaan merkittyjen käyttäjien läsnäolotietoja ja ilmoittaa muutoksista.

Läsnäololle ei ole kehitetty standardia, vaan kaikki ohjelmistovalmistajat käyttävät ei-standardeja protokollia ohjelmistoissaan. Tällöin eri sovellusten välinen keskinäinen kommunikointi on mahdotonta. Hyvänä esimerkkinä tästä ovat AOL Instant Messenger, Yahoo! Messenger ja MSN Messenger, joista millään ei pysty kommunikoimaan toisten ohjelmien käyttäjien kanssa. ”Instant Messaging and Presence”-työryhmä (IMPP Working Group) on perustettu kehittämään standardiprotokolla, jotta ohjelmistojen kehittäjät voivat luoda yhteensopivia ohjelmistoja.

Langattomien laitteiden määrän arvoidaan kasvavan tulevina vuosina huomattavasti, eikä ole järkevää olettaa, että eri protokollat ovat käytössä langattomissa laitteissa. Lisäksi langaton infrastruktuuri on kypsymässä nopeasti käyttökelpoiseksi ympäristöksi. Tämän vuoksi IMPP-työryhmän on otettava huomioon myös lähitulevaisuus suunnitellessaan standardia. Protokollan käyttöympäristö on tulevaisuudessa todennäköisesti mobiili pääte-laite, jolloin pitää huomioida korkea saantiviive, pieni kaistanleveys ja mahdollisesti jaksottainen yhteydessäolo, vaatimaton laskentateho, virtalähteen rajoitukset, pienet näytöt jne.

5.1. Arkkitehtuuri

Läsnäolotietopalvelu (presence service) vastaanottaa, tallentaa ja jakelee läsnäolotietoa. Läsnäolotietopalvelun abstrakti kuvaus on määritelty dokumentissa ”RFC 2778 – A Model for Presence and Instant Messaging”. Kaikki palvelun käsittelemä tieto on läsnäolotietoa. Läsnäolotietopalvelu voidaan ajatella laajennettuna versiona aikaisemmin esitetystä paikkatietopalvelusta. Läsnäolotietopalvelu ei korvaa paikkatietopalvelua, mutta se laajentaa sen käyttämään myös läsnäolotietoa.

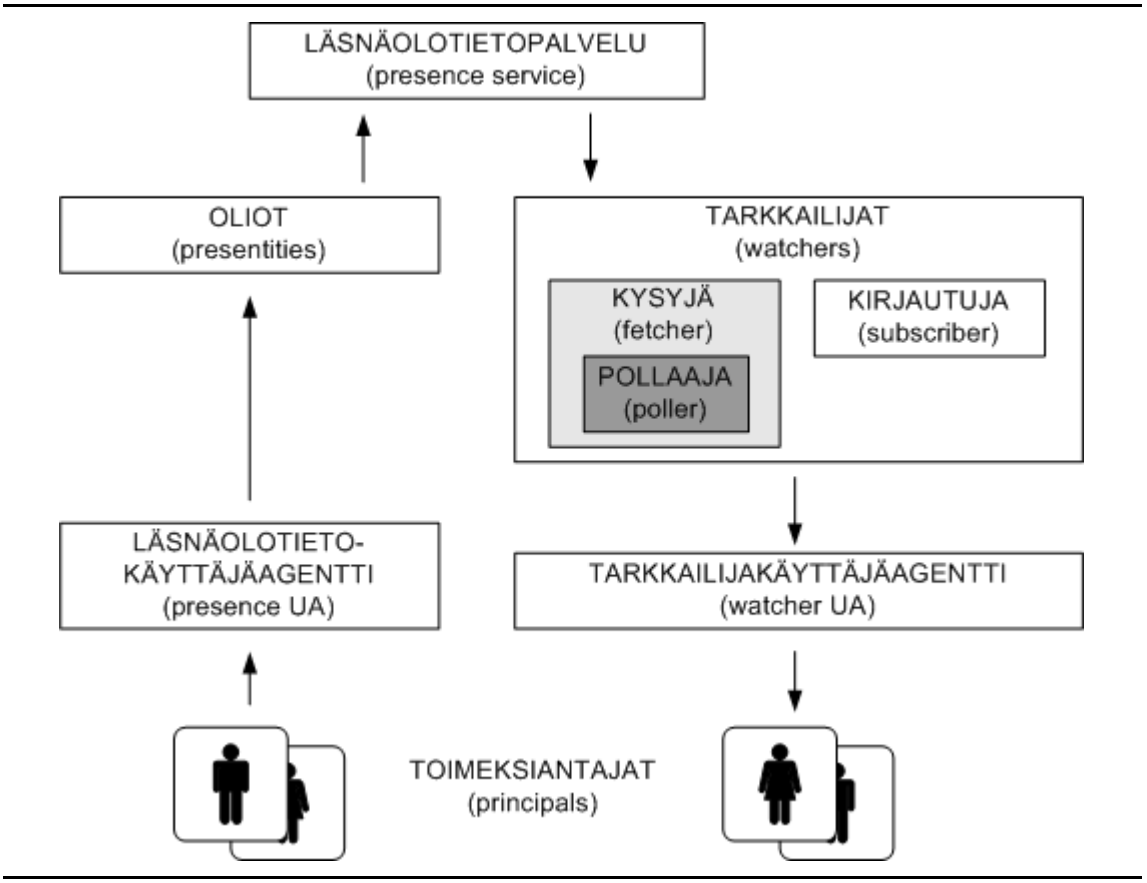
Läsnäolotietopalvelu voidaan jakaa kahteen eri osaan, joilla molemmilla on oma tehtävänsä. Usein nämä osat ovat yhdistetty toteutettaessa läsnäolotietopalvelu, mutta ne ovat loogisesti kaksi eri asiaa, joten myös niitä käsitellään tässä erikseen.

Ensimmäisen osan muodostavat ”oliot” (presentity). Ne tuottavat läsnäolotietoja jaettavaksi ja tallennettavaksi läsnäolotietopalvelulle eli generoivat NOTIFY-viestejä. Olio voi olla mikä tahansa asia, joka tuottaa läsnäolotietoa läsnäolopalvelulle. Yleensä nämä ovat SIP-käyttäjiä, jotka voidaan tunnistaa SIP-osoittein [Liscano].

Toisen osan muodostavat tarkkailijat (watcher). Niiden tehtävä on pyytää olioiden läsnäolotietoja. Myös nämä ovat yleensä SIP-käyttäjiä, jotka voidaan tunnistaa SIP-osoittein. Asia on havainnollistettu kuvassa 12 [Koskela].

Tarkkailijat jakaantuvat lisäksi kahteen eri palveluun, kysyjä- (fetcher) ja kirjautujapalveluun (subscriber). Kysyjäpalvelu tiedustelee tietyn käyttäjän nykyistä tilaa läsnäolotietopalvelulta epäsäännöllisesti. Pollaaja (poller) on kysyjäpalvelu, joka tiedustelee läsnäolotietoja säännöllisesti läsnäolotietopalvelulta. Kirjautujapalvelu kuuntelee läsnäolotietopalvelun ilmoituksia tietyn käyttäjän läsnäolotietojen sen hetkisistä tai tulevista muutoksista [Koskela].

Läsnäolotietojärjestelmiä käyttäviä ihmisiä, ryhmiä tai ohjelmistoja kutsutaan nimellä toimeksiantajat (principals). Toimeksiantajat keskustelevat läsnäolotietopalvelun kanssa käyttäjäagenttien välityksellä. Käyttäjäagentit tarjoavat pelkästään liitynnän toimeksiantajan sekä järjestelmän jonkin palvelun kanssa. Usein useita käyttäjäagentteja on toteutettu samaan päätelaitteeseen. Esimerkkejä käyttäjäagenteista ovat “inbox user agent”, ”sender user agent”, “presence user agent” ja “watcher user agent”. On täysin sovelluskohtaista miten tämä kommunikointi tapahtuu. Läsnäolotietopalvelu tietää vain, että kaksi toimeksiantajaa joko ovat tai eivät ole samat [RFC 2778].



Kuva 12. Yleiskuva läsnäolotietopalvelusta

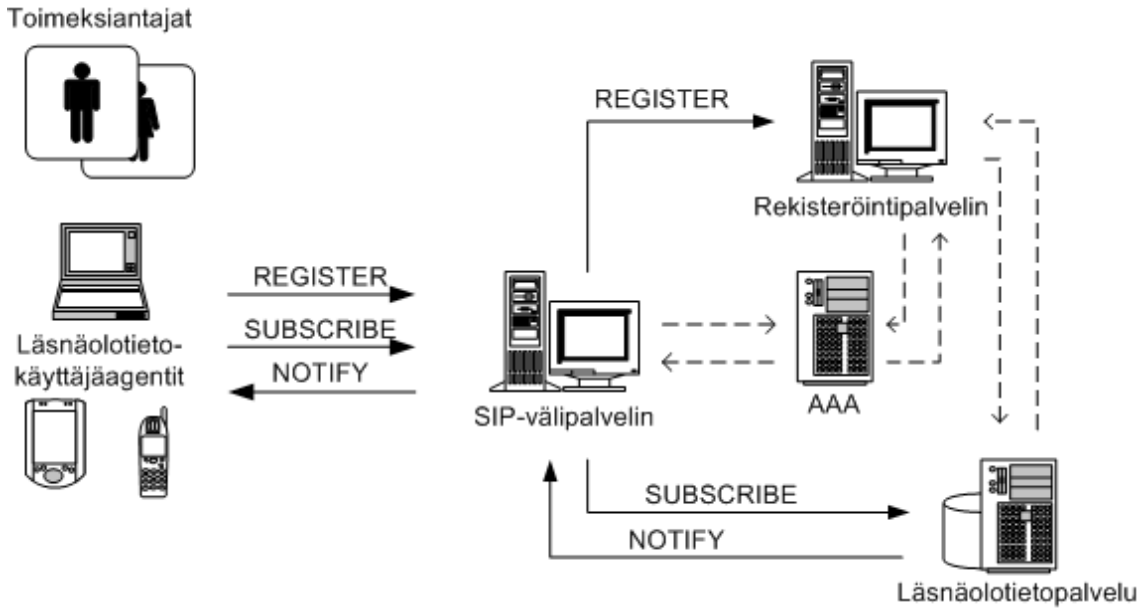
Läsnäolotietokäyttäjäagentti (Presence User Agent - PUA) muokkaa olion läsnäolotietoja. Läsnäolotietojen muokkaaminen voi johtua esimerkiksi SIP REGISTER –viestin lähettämisestä lisättäessä uutta sidontaa. Oliolla voi olla useita läsnäolotietokäyttäjäagentteja. Tämä tarkoittaa, että käyttäjällä voi olla useita eri päätelaitteita (kuten matkapuhelin ja kämmen-mikro), joista jokainen itsenäisesti luo läsnäolotietoa. Läsnäolotietokäyttäjäagentit työntävät (push) tietoa läsnäolojärjestelmään, mutta ne eivät vastaanota SUBSCRIBE-viestejä, eivätkä lähetä NOTIFY-viestejä [Rosenberg 2].

Läsnäoloagentti (Presence Agent - PA) on SIP-käyttäjäagentti, joka vastaanottaa SUBSCRIBE-viestejä, vastaa niihin ja luo ilmoituksia (notifications) läsnäolotitilan muutoksista. Läsnäoloagentilla täytyy olla aina tieto olion sen hetkisestä tilasta. Tämä tarkoittaa, että sillä tulee olla pääsy läsnäolotietoihin, joita läsnäolokäyttäjäagentit muokkaavat. Yksi tapa toteuttaa tämä on yhdistää läsnäoloagentti välipalvelimen tai rekisterinpitäjäpalvelimen kanssa. Toinen toteutustapa on yhdistää läsnäoloagentti läsnäolotietokäyttäjäagentin kanssa. Läsnäoloagentti on aina osoitettavissa SIP Internet-resurssin tunnisteella, joka yksilöi käyttäjän (esimerkiksi sip:petrilintula@uta.fi). Käyttäjällä voi olla useita läsnäoloagentteja, joista

jokainen vastaa tietystä osasta kirjautumisia, jotka käyttäjällä on sillä hetkellä tehtyinä [Rosenberg 2].

Läsnäolotietopalvelin (presence server) on fyysinen laite. Se toimii joko läsnäolotietoagenttina tai välipalvelimena kirjautumispyynnöille. Toimiessaan läsnäolotietoagenttina, se on myös tietoinen käyttäjän läsnäolotiedosta. Toimiessaan välipalvelimena, kirjautumispyynnöt ohjataan toiselle instanssille, joka toimii läsnäolotietoagenttina.

Kuvassa 13 on havainnollistettu yksinkertainen läsnäolotietojärjestelmä. Käyttäjän läsnäolotietoagentti lähettää järjestelmään REGISTER-viestin ja ilmoittaa, mistä käyttäjä on tavoitettavissa. Välipalvelin tarkistaa, onko käyttäjä oikeutettu käyttämään verkkoa ja välittää REGISTER-viestin rekisterinpitäjä-palvelimelle. Rekisterinpitäjä tarkistaa, onko käyttäjä oikeutettu tekemään rekisteröintejä ja tallentaa rekisteröinnin läsnäolotietopalveluun. Muut käyttäjät voivat rekisteröityä kuuntelemaan käyttäjän läsnäolotietoja, nämä sidonnat tallennetaan läsnäolotietopalveluun. Läsnäolotietopalvelu luo NOTIFY-viestejä, jotka lähetetään käyttäjän läsnäolotietoja kuuntelemaan rekisteröityneille käyttäjille.



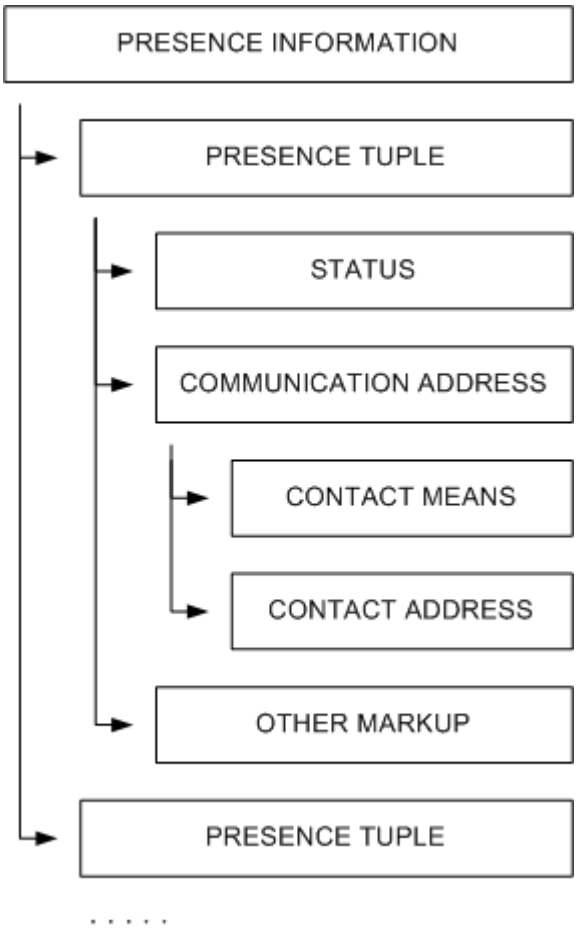
Kuva 13. Läsnäolotietojärjestelmä

5.2. Läsnäolotiedon formaatti

Läsnäolotiedon formaatti on nimeltään "Presence Information Data Format" (PIDF) ja sen rakenne on esitetty kuvassa 14. PIDF on XML-pohjainen (eXtensible Markup Language) eli siten tekstipohjainen. PIDF määrittelee minimivaatimukset formaatille ja sitä voidaan laajentaa käsittämään lisää elementtejä.

Läsnäolotieto koostuu vaihtelevasta määrästä elementtejä, joita kutsutaan nimellä ”presence tuple”. Jokainen ”presence tuple”-elementti sisältää status-merkin, joka voi olla esimerkiksi online, offline, kiireinen tai poissa, valinnaisen kommunikointiosoitteen ja valinnaisen ”other presence markup”-elementin. Kommunikointitosoite-elementti sisältää kommunikointitapa- ja yhteystieto-elementin. Esimerkkejä kommunikointiosoite-elementin sisällöstä ovat suoraviestintäosoite ja puhelinnumero.

Statuksella on aina vähintään kaksi tilaa, jotka toimivat yhdessa suora-viestinnän kanssa: avoin, jolloin suoraviestit hyväksytään, ja suljettu, jolloin suoraviestejä ei hyväksytä. Nämä tilat voivat toimia myös muiden kommunkiointitapojen, ei pelkästään suoraviestinnän kanssa.



Kuva 14. Läsnäolotiedon formaatti

Koodiesimerkissä 13 on esitetty läsnäolotietoviesti, joka sisältää kaksi eri läsnäolotietoa. Ensimmäinen ”tuple” ilmoittaa käyttäjän tilan tällä hetkellä ja toinen ilmoittaa käyttäjän tilan ensi viikolla.

```
<?xml version="1.0" encoding="UTF-8"?>
  <presence xmlns="urn:ietf:params:xml:ns:pidf"
      xmlns:im="urn:ietf:params:xml:ns:pidf:im"
      xmlns:myex="http://id.example.com/presence/"
      entity="pres:someone@example.com">
    <tuple id="bs35r9">
      <status>
        <basic>open</basic>
        <im:im>busy</im:im>
        <myex:location>home</myex:location>
      </status>
      <contact priority="0.8">im:henkilo@mobilecarrier.net</contact>
      <note xml:lang="en">Don't Disturb Please!</note>
      <note xml:lang="fi">Ala hairitse, kiitos.</note>
      <timestamp>2001-10-27T16:49:29Z</timestamp>
    </tuple>
    <tuple id="eg92n8">
      <status>
        <basic>open</basic>
      </status>
      <contact priority="1.0">mailto:someone@example.com</contact>
    </tuple>
    <note>I'll be in Tokyo next week</note>
  </presence>
```

Koodiesimerkki 13. Läsnäolotiedon formaatti

5.3. Toiminnot

5.3.1. Rekisteröityminen

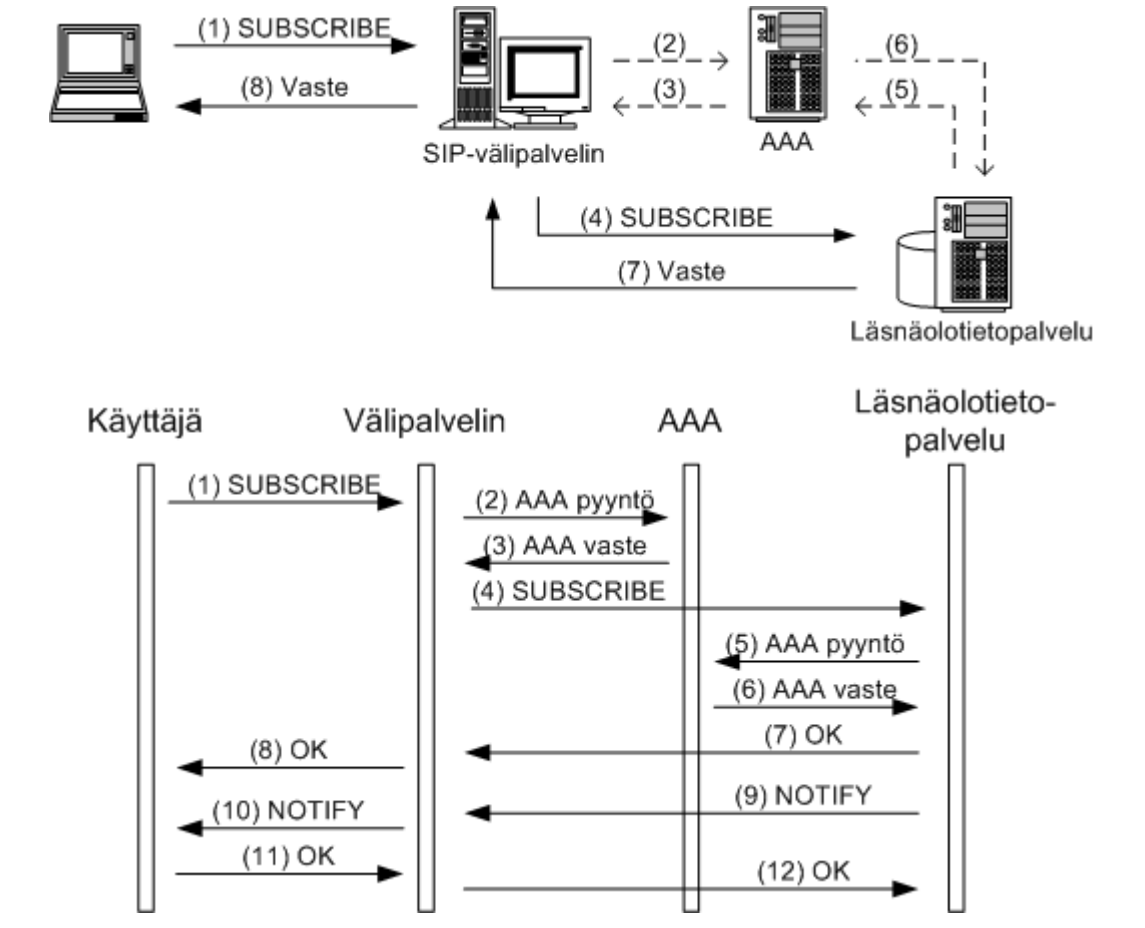
Käyttäjä rekisteröityy läsnäolotietojärjestelmään REGISTER-viestillä, jota käsiteltiin tarkemmin kohdassa 3.7. Yhdelle päätelaitteelle (yhteen ip-osoitteeseen) voi rekisteröityä useita käyttäjiä. Rekisteröityminen on erittäin tärkeää, koska muuten järjestelmä ei tiedä, mistä käyttäjän tavoittaa. Se on koko läsnäolotietojärjestelmän perusehto.

Liikkuvuus (mobility) toteutetaan SIP-ympäristössä samankaltaisesti kuin matkapuhelimissakin. Päätelaite lähettää jatkuvasti REGISTER-viestejä, joiden avulla verkolla ja palvelimilla on tieto, mihin ottaa yhteyttä, ja missä päätelaite kulloinkin sijaitsee. Liikkuvuus tarkoittaa myös paikkatietojen päivitystä, ei

pelkästään puheyhteydellistä liikkuvuutta. Esimerkiksi alkuperäinen rekisteröintitieto voi olla puhelintieto ja uusi rekisteröintitieto vaikka sähköposti ja puhepostilaatikko (voice mail box).

5.3.2. Kirjautuminen

Käyttäjän halutessa saada tietoja toisen käyttäjän läsnäolotiedoista, hän lähettää järjestelmään SUBSCRIBE-viestin (kuva 15). Viesti kulkee välipalvelimien kautta läsnäolotietopalvelulle, jossa käyttäjä lisätään kuuntelemaan halutun käyttäjän läsnäolotietoja. Molemmat välipalvelin ja läsnäolotietopalvelu tarkistavat AAA-palvelimelta, onko käyttäjä oikeutettu tekemään toiminnon. Järjestelmä vastaa SUBSCRIBE-viestiin SIP OK –viestillä ilmoittaen hyväksyneensä pyynnön. Tämän jälkeen järjestelmä lähettää käyttäjälle NOTIFY-viestin, joka sisältää kohteen tämänhetkiset läsnäolotiedot. Koodi-esimerkissä 14 on esitetty järjestelmän lähettämä SUBSCRIBE-viesti sekä koodiesimerkissä 15 järjestelmän SIP OK-vaste pyyntöön.



Kuva 15. Subscribe-pyynnön luonti

```
SUBSCRIBE sip:resource@example.com SIP/2.0
Via: SIP/2.0/TCP watcherhost.example.com;branch=z9hG4bKnashds7
To: <sip:resource@example.com>
From: <sip:user@example.com>;tag=xfg9
Call-ID: 2010@watcherhost.example.com
CSeq: 17766 SUBSCRIBE
Max-Forwards: 70
Event: presence
Accept: application/cpim-pidf+xml
Contact: <sip:user@watcherhost.example.com>
Expires: 600
Content-Length: 0
```

Koodiesimerkki 14. (1) SUBSCRIBE-viesti

```
SIP/2.0 200 OK
Via: SIP/2.0/TCP watcherhost.example.com;branch=z9hG4bKnashds7
    ;received=192.0.2.1
To: <sip:resource@example.com>;tag=ffd2
From: <sip:user@example.com>;tag=xfg9
Call-ID: 2010@watcherhost.example.com
CSeq: 17766 SUBSCRIBE
Event: presence
Expires: 600
Contact: sip:server.example.com
Content-Length: 0
```

Koodiesimerkki 15. (7) OK-viesti

SUBSCRIBE-pyynnössä määritellään sen voimassaoloaika. Pyynnön aika-arvo ollessa 0, se on yksittäinen pyyntö, eikä käyttäjä halua jäädä kuuntelemaan läsnäolotietoja. Jokaiselle SUBSCRIBE-pyynnölle luodaan yksilöllinen ”SubscriptID”, josta sen jatkossa tunnistaa.

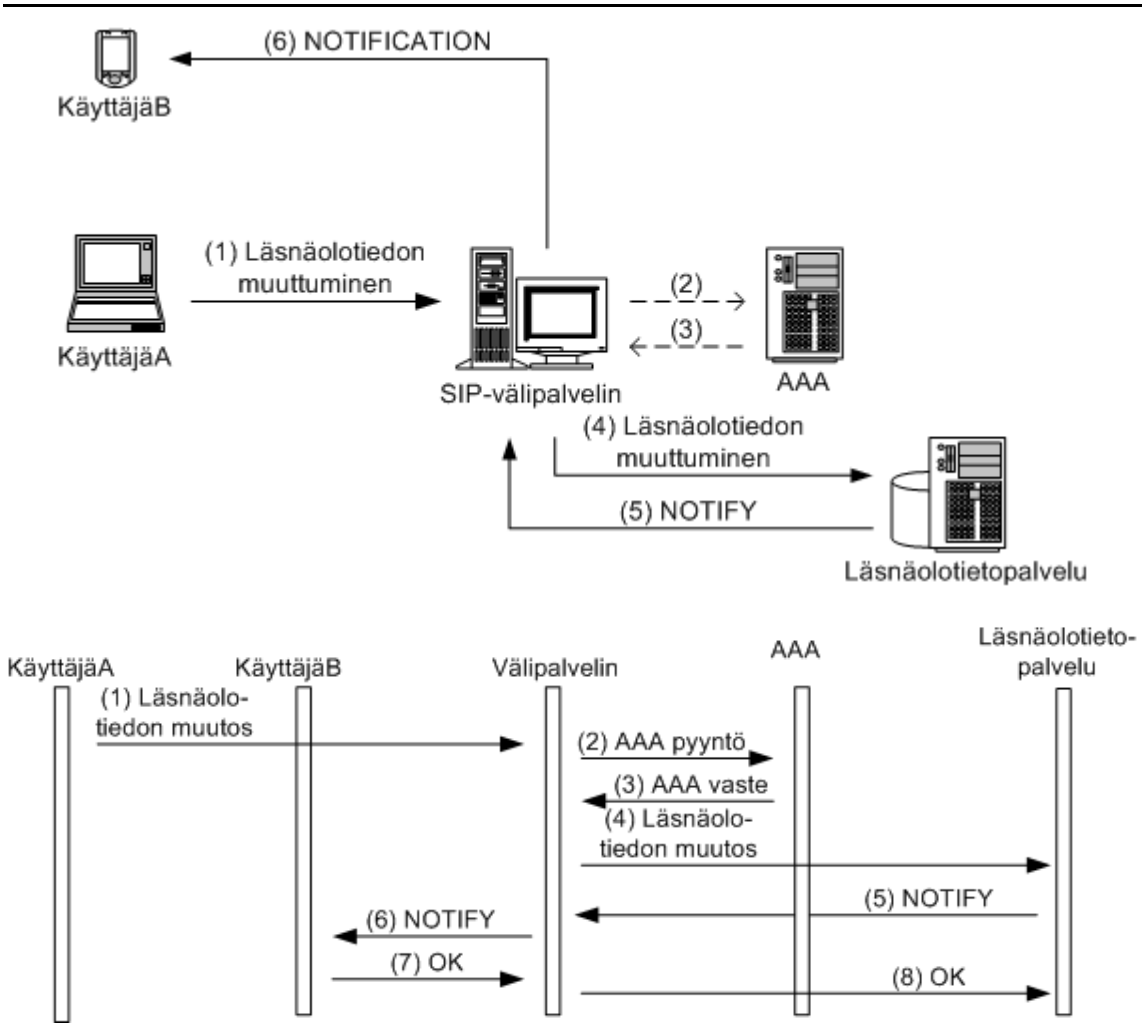
Käyttäjän halutessa lopettaa läsnäolotietojen tarkkailun, se lähettää järjestelmälle SUBSCRIBE-viestin, jonka aika-arvo on 0. Läsnäolotietopalvelu tarkistaa, löytyykö viestissä oleva ”SubscriptID” palvelusta ja mikäli löytyy, se poistaa poistaa käyttäjän tarkkailemasta toisen käyttäjän läsnäolotietoja.

5.3.3. Notify

Järjestelmä luo NOTIFY-viestin käyttäjän läsnäolotietojen muuttuessa. NOTIFY-viesti lähetetään kaikille toisen käyttäjän läsnäolotietoja

kuuntelemaan kirjautuneille käyttäjille. NOTIFY-viestiin käyttäjät vastaavat OK-viestillä.

Kuvassa 16 KäyttäjäA:n läsnäolotiedot muuttuvat ja hänen läsnäolotietoagenttinsa luo automaattisesti ilmoituksen järjestelmään. Välipalvelin tarkistaa, että käyttäjä on oikeutettu käyttämään verkkoa ja välittää läsnäolotietojen muutoksen läsnäolotietopalvelulle. Läsnäolotietopalvelu tallentaa muutokset ja lähettää KäyttäjäA:n tietoja kuuntelemaan kirjautuneelle KäyttäjäB:lle NOTIFY-viestin (koodiesimerkki 16). KäyttäjäB vastaa NOTIFY-viestiin OK-viestillä ilmoittaen saaneensa tiedon läsnäolotietojen muutoksesta.



Kuva 16. Käyttäjä päivittää läsnäolotietoaan

```
NOTIFY sip:user@watcherhost.example.com SIP/2.0
Via: SIP/2.0/TCP server.example.com;branch=z9hG4bKna998sl
From: <sip:resource@example.com>;tag=ffd2
To: <sip:user@example.com>;tag=xfg9
Call-ID: 2010@watcherhost.example.com
CSeq: 8776 NOTIFY
Event: presence
Subscription-State: active;expires=543
Max-Forwards: 70
Contact: sip:server.example.com
Content-Type: application/cpim-pidf+xml
Content-Length: ...

[uusi PIDF dokumentti]
```

Koodiesimerkki 16. (6) NOTIFY-viesti

6. Tietoturva

Tietoturva on oleellinen asia, koska tietoa lähetetään vihamielisenä ympäristönä pidetyssä Internetissä. Tietoturva SIP:issä rajoittuu istunnon luontiin ja sen muokkaamiseen kuten itse SIP:kin. Tämän vuoksi istunto voidaan perustaa turvallisesti, mutta istunnossa lähetettävä data voi olla salaamatonta. Tällöin data voidaan helposti kaapata ja tulkita.

Tietoturva ei ole yksittäinen laaja ominaisuus järjestelmässä, vaan ennemminkin kokoelma erillisiä ja itsenäisiä ominaisuuksia. Suurimmat tietoturvauhat ovat:

1. Valtuuttamaton SIP / SIMPLE -viestintä tai läsnäolopalveluiden käyttö, joko uutena käyttäjänä tai olemassa olevana käyttäjänä,
2. Vihamielinen tai vieras käyttäjä tai palvelu esiintyy autenttisena verkko osana loppukäyttäjälle,
3. Salakuuntelu (eavesdropping) ja viestien muuntelu,
4. Valtuuttamaton pääsy lukemaan tai muokkaaminen toisen käyttäjän tietoja tai toisen käyttäjän tallentamia viestejä.

6.1. Käyttäjien autentikointi

Palvelinten on pystyttävä autentikoimaan (authenticate) käyttäjät ja käyttäjien on pystyttävä autentikoimaan palvelimet. Autentikointi tarkoittaa varmistumista toisen osapuolen identiteetistä. Se ei ole sama asia kuin lupa käyttää verkon palveluita. Käyttäjä voidaan autentikoida, mutta häneltä voidaan siitä huolimatta evätä palvelun käyttö. Palvelin tai käyttäjän päätelaite voi milloin tahansa haastaa toisen osapuolen todistamaan henkilöllisyytensä. Tunnistamisen jälkeen riippuu osapuolista, onko heillä riittävästi oikeuksia keskustella keskenään

SIP perustuu HTTP:hen, joten se voi käyttää HTTP:n haastepohjaista (challenge-based) koostetunnistetta (digest authentication) autentikointimekanisminään. HTTP:ssä on myös perustunnistusmekanismi (basic authentication), jota ei pidä käyttää sen heikon tietoturvan takia. [RFC 2617].

Perustunnistuksessa asiakas lähettää käyttäjätunnuksen ja salasanan valtuustietonaan (creditentials). Käyttäjätunnus ja salasana lähetetään tekstimuodossa, jonka takia tunnistusta ei pidä käyttää. Koostetunnistuksessa näin ei tapahdu, vaikka sekin perustuu käyttäjätunnus- ja salasanamekanismiin. Asiakas lähettää palvelimelle käyttäjätunnuksesta ja salasanasta lasketun tarkastussumman, jonka perusteella palvelin tietää asiakkaan tuntevan oikean

salasanan. Käyttäjätunnusta ja salasanaa ei lähetetä ollenkaan asiakkaalta palvelimelle. Koostetunnistus kuitenkin tarjoaa vain autentikointi-mekanismin, ei mekanismia varmistaa viestin eheyttä eikä luottamuksellisuutta [RFC 2617].

Eräs merkittävimistä rajoituksista koostetunnisteen käytölle SIP:issä on, ettei eheysmekanismi (integrity mechanism) toimi kovinkaan hyvin. Se tarjoaa tietoturvaa Internet-resussin tunnisteelle sekä viestille, muttei lainkaan otsakekentille, jotka käyttäjä varmasti haluaisi myös suojata.

Toinen rajoite koostetunnisteen käytölle on alueiden (realms) laajuus. Koostetunniste on hyvä menetelmä, kun käyttäjä haluaa autentikoida itsensä käyttämään resurssia, jonne hänellä on olemassaoleva liityntä (esimerkiksi kuten palveluntarjoajalle jonka asiakas käyttäjä on). Autentikoitaessa toisiin alueisiin, käyttäjä saattaa kohdata hankaluuksia.

6.2. Viestin eheys ja luottamuksellisuus

Viestin eheys (integrity) tarkoittaa varmuutta, ettei viestiä ole muutettu matkalla palvelimelta asiakkaalle. Viestin luottamuksellisuus (confidentiality) vastaavasti tarkoittaa, että viestin pystyy lukemaan vain vastaanottaja jolle se on tarkoitettu.

Yleinen mekanismi viestin eheyden ja luottamuksellisuuden varmista-miseksi on Secure/Multipurpose Internet Mail Extension (S/MIME) [RFC 2633]. S/MIME:n avulla viestin sisältö allekirjoitetaan käyttäen salaisen ja julkisen avaimen tekniikkaa. Käyttäjä tai sovellus laskee digitaalisen allekirjoituksen käyttäen salaista avainta ja viestin sisältöä. Muut käyttäjät pystyvät avaamaan viestin julkisella avaimella ja varmistamaan, että viesti on alkuperäisessä muodossa. Vastaavasti muut käyttäjät voivat salata takaisin lähetettävän viestin julkisella avaimella, jonka vain salaisen avaimen omaava käyttäjä pystyy avaamaan.

6.3. Yksityisyys

SIP-viestit saattavat usein sisältää henkilökohtaista tai arkaluonteista informaatiota niiden lähettäjistä. Ei pelkästään se, mitä viestissä sanotaan, mutta myös kenen kanssa kommunikoidaan, milloin, miten pitkään ja mistä istuntoon osallistutaan. Monet sovellukset ja käyttäjät vaativat, että tämän-kaltainen tieto pitää piilottaa tahoilta, joiden sitä ei tarvitse tietää.

On olemassa myös epäsuorempia keinoja, jolla henkilökohtainen informaatio voidaan selvittää. Jos käyttäjä tai palvelu on tavoitettavissa arvattavasta osoitteesta, perinteinen metodi omistaa salattu puhelinnumero ei enää toimi. Läsnäolotietopalvelu voi loukata vastaanottajan yksityisyyttä paljastamalla hänen sijaintinsa yhteydenottajalle. Luonnollisesti on enem-

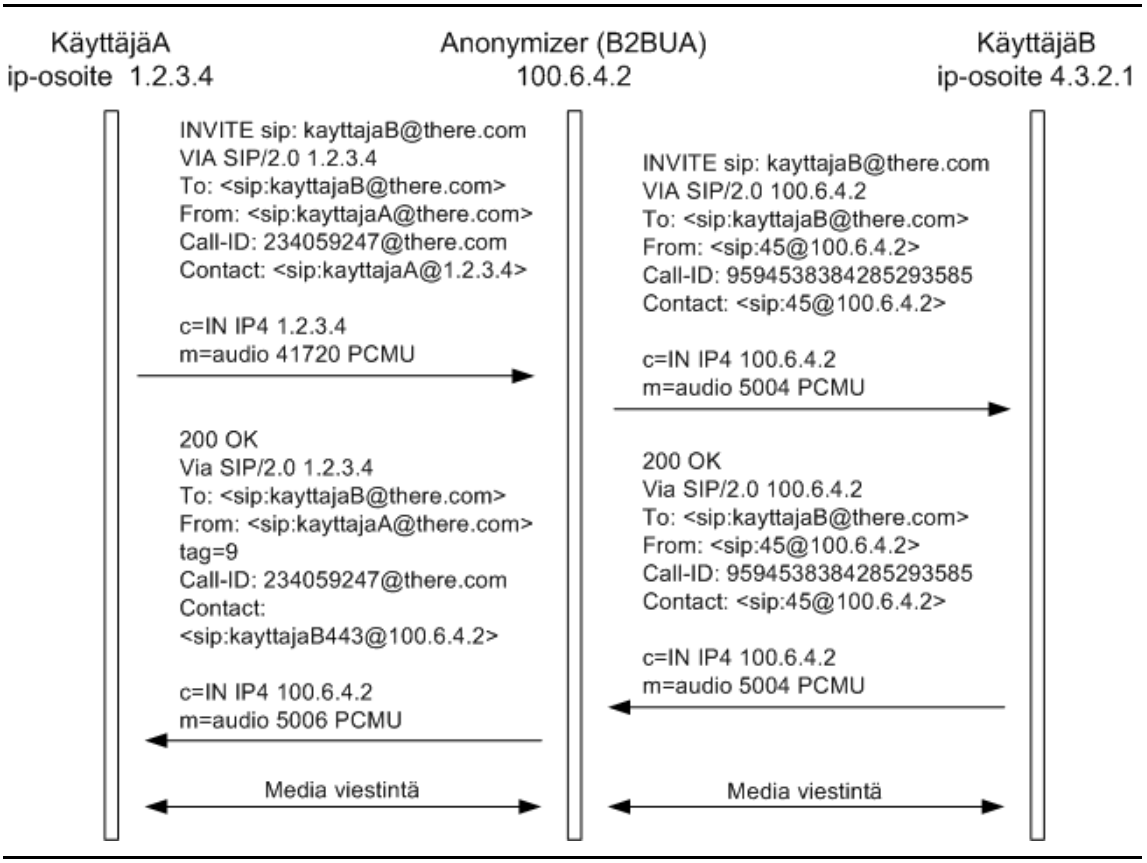
mänkin läsnäolotietopalvelun ongelma, minkälaista tietoa käyttäjistään se ulkopuolelle antaa.

Käyttäjiä voidaan ryhmitellä ja tietyille ryhmille voidaan näyttää enemmän tietoa kuin toisille ryhmille. Tätä ongelmaa ei vielä ole käsitelty paljoakaan. Toiset käyttäjät haluavat antaa itsestään paljon tietoa, kun taas toiset eivät halua antaa itsestään mitään tietoa ulospäin.

Käyttäjät saattavat haluta kätkeä henkilökohtaista tietoa otsaketiedoista, joista voisi päätellä käyttäjän identiteetin. Tämä ei koske pelkästään ”From”-kenttää ja viestin alkuperää koskevia kenttiä, mutta myös ”To”-kenttää. Ehkä halutaan kätkeä kutsumanimi, jolla toista ihmistä kutsutaan omassa päätelaitteessa.

Myös läsnäolopalvelussa yksityisyys saattaa aiheutta ongelmia. Henkilö ei välttämättä halua paljastaa olevansa tilaajana toisen käyttäjän läsnäolotiedoille. Lisäksi NOTIFY-viestit voivat sisältää arkaluontoista informaatiota käyttäjästä.

Yksityisyyden suojaamiseksi SIP-verkon on mahdollista käyttää back-to-back käyttäjäagenttia anonyymipalvelun toteuttamiseksi. Palvelussa käyttäjän ip-osoite, URL tai muut identifiointitiedot voidaan estää näkymästä vastaanottavalle osapuolelle. Tämä tapahtuu muodostamalla kaksi täysin erillistä istuntoa B2BUA:n välittäessä signalointi- ja mediainformaation käyttäjältä toiselle. Kumpikin osapuoli lähettää SIP- ja RTP- paketit B2BUA:lle, eivätkä toisilleen. Kun puhelu on lopetettu, anonymointipalvelu poistaa kaikki loki- ja tilatiedot. Esimerkki tälläisestä B2BUA:n käytöstä on esitetty kuvassa 17. Kuvassa on myös esitetty viestit, jotka tapahtumassa välitetään. Vertailtaessa viestien lähettäjien ja vastaanottajien osoitteita nähdään miten anonyymipalvelu kätkee toisen osapuolen identifiointitiedot [Internet communication using SIP].



Kuva 17. B2BUA:n käyttö anonyymiteetin luomiseksi

6.4. Tietoturvamekanismit

SIP ei ole yksinkertainen tai helppo protokolla turvata, koska

- se käyttää useita välipalvelimia,
- sillä on monimutkaiset luottosuhteet,
- se saattaa käyttää viestintään verkkoelementtejä, joihin ei voi ollenkaan luottaa,
- sen käyttäjältä-käyttäjälle –operaatiot ovat vaikeita suojata.

Näihin haasteisiin vastaaminen vaatii tietoturvaratkaisuja, jotka pystyvät toimimaan itsenäisesti ja useissa eri ympäristöissä ja käyttötarkoituksissa. Ei ole olemassa yksittäistä ratkaisua, vaan näiden tarpeiden tyydyttämiseksi tarvitaan useita eri toimintoihin ja kohteisiin soveltuvia erillisiä mekanismeja. Hyvänä puolena on, että mikä tahansa istuntoon liitetty media voidaan salata loppukäyttäjältä loppukäyttäjälle itsenäisesti SIP:istä riippumatta. Kaikilla tietoturvamekanismeillä on kuitenkin rajoituksensa, jotka tulee ymmärtää ennen kuin niitä voi käyttää oikein ja turvallisesti.

Uusien tietoturvamekanismien määrittämisen sijaan SIP käyttää mahdollisimman paljon olemassa olevia tietoturva ratkaisuja, joita on otettu HTTP ja SMTP:stä.

Viestien täydellinen salaus tarjoaa aina parhaan mahdollisuuden säilyttää viestinnän luottamuksellisuus - se myös takaa, ettei viestiä ole muokattu matkalla. SIP-pyyntöjä ja –vasteita ei kuitenkaan voida salata loppukäyttäjältä loppukäyttäjälle kokonaisuudessaan, koska viesti itsessään sisältää sen perille menemiseksi välttämättömiä tietoja kuten ”request-URI”-, ”Route”- ja ”Via”-kentät. Salattaessa nämä kentät, välipalvelimet eivät osaisi välittää viestejä oikeisiin osoitteisiin. Välipalvelinten tulee pystyä muokkaamaan viestejä, kuten lisäämällä ”Via”-kentän arvoja. SIP-käyttäjäagenttien tulee ainakin jossain määrin luottaa välipalvelimiin. Alemman tason tietoturvamekanismeja suositellaankin käytettäväksi. Ne mahdollistavat koko SIP-viestin salaamisen kahden laitteen (SIP-palvelimen) välillä. Tälläistä salausta kutsutaan hyppyjen väliseksi (hop-by-hop) salaukseksi. Niiden avulla loppukäyttäjien on mahdollista varmistaa käyttämiensä välipalvelinten aitous ja turvallisuus. Näin koko SIP-viesti on mahdollista salata loppukäyttäjältä loppukäyttäjälle [RFC 3261].

Kuljetus- tai verkkokerroksen tietoturva salaa signalointiliikenteen varmistaen viestin luottamuksellisuuden ja eheyden. Sertifikaatteja käytetään myös luotaessa alemman tason tietoturvaa. Lisäksi sertifikaatteja voidaan käyttää tarjoamaan keinot autentikoida monissa arkkitehtuureissa [RFC 3261].

Kaksi suosittua vaihtoehtoa tietoturvan tarjoamiseksi kuljetus- ja verkkokerroksissa ovat IPSec ja TLS. IPSec on kokoelma verkkokerroksen protokolla-työkaluja, joita voidaan käyttää luomaan turvallinen yhteys perinteisen IP:een sijaan. IPSec:iä käytetään arkkitehtuureissa, joissa on joukko toisiinsa luottavia isäntäpalvelimia tai toimialueita. Se yleensä toteutetaan käyttöjärjestelmätasolla isäntäpalvelimeen tai ”security gateway”:hin, joka tarjoaa luottamuksellisuutta ja eheyttä kaikelle sen kautta tietyistä liitynnöistä kulkevalle liikenteelle (kuten VPN-arkkitehtuurissa). IPSec:iä voidaan myös käyttää hyppyjen välillä [RFC 3261].

TLS tarjoaa tietoturvaa kuljetuskerroksen liityntäpohjaisille protokollille, kuten TCP). Se voidaan määrittää halutuksi kuljetusprotokollaksi ”Via”-kentän avulla. TLS soveltuu tarvittaessa hyppyjen välistä tietoturvaa isäntäpalvelinten välillä, joilla ei ole ennalta olemassa luottamussuhdetta toisiinsa. Esimerkiksi Alice luottaa hänen välipalvelimeensa, joka puolestaan sertifikaattivaihdon jälkeen päättää luottaa Bobin välipalvelimeen, johon Bob luottaa. Näin Bob ja Alice voivat kommunikoida turvallisesti. TLS:n suurin ongelma on, ettei sitä voi käyttää UDP:n kanssa. TLS vaatii liityntäpohjaisen siirtoprotokollan, joka tässä tapauksessa tarkoittaa TCP:tä. TLS:n käyttö jokaisessa elementissä viestien reitillä varmistaa, että se on kaikkialla käytössä [RFC 3261].

Loppukäyttäjän on tärkeää tiedostaa kommunikoivansa oikean palvelu-elementin kanssa. Sekä TLS että IPSec tarjoavat mekanismeja, joiden avulla loppukäyttäjän on mahdollista varmistaa ensimmäisen palveluelementin identiteetti palveluelementtien ketjussa. Loppukäyttäjän pitää kuitenkin yleensä luottaa muihin luodakseen turvallisen yhteyden. Esimerkki tästä on SIPS Internet-resurssin käyttö.

SIP-viestit sisältävät MIME-runkoja (bodies) ja MIME-standardi sisältää mekanismit suojaamaan (securing) MIME-sisällön eheyttä ja luottamuksellisuutta. Tässä on laskettu näihin kuuluvaksi ”multipart/signed” ja ”application/pkcs7-mime” MIME-tyypit (rfc1847, rfc2630, rfc2633). PGP-mekanismia viestien otsake-kenttien (header) sekä viestien runkojen salaamiseksi ei enää suositella käytettäväksi [RFC 3261].

Kuten aiemmin todettiin koko SIP-viestin salaus ei ole mahdollista. S/MIME kuitenkin mahdollistaa SIP-käyttäjäagenttien salata MIME rungot. Se tarjoaa luottamuksellisuutta loppukäyttäjältä loppukäyttäjälle, viestien runkojen eheyttä ja authentikointia. S/MIME:ä on myös mahdollista käyttää luomaan jonkin näköistä eheyttä ja luottamuksellisuutta SIP-otsakekentille käyttämällä tunnelointia [RFC 3261].

S/MIME:n suurin puute on sen loppukäyttäjien julkisten avainten infrastruktuurin puute. Käytettäessä itseluotuja sertifikaatteja (tai sertifikaatteja joita ei voida varmentaa), SIP-pohjainen avaintenvaihto on altis ”man-in-the-middle”-hyökkäykselle, jolloin hyökkääjä voi mahdollisesti lukea ja muuttaa viestien sisältöä.

Vastaanottajien omaatessa toistensa sertifikaatit, S/MIME mahdollistaa salattujen viestien välityksen ilman riskiä. On mahdollista, että kyseistä sertifikaattia ei enää ole ip-osoitteessa olevassa laitteessa, jota käyttäjä aikaisemmin käytti. Näin hän ei pysty avaamaan salattuja viestejä, joka johtaa virhetilanteisiin käyttäjien välillä. Erityisesti tilanne ilmenee silloin, kun salattu viesti on jaettu (forked). Avaimet ovat erityisen käytännöllisiä liitettäessä S/MIME:een ne henkilön osoitteeseen (petri.lintula@uta.fi) eikä laitteeseen, ne ovat erityisen käytännöllisiä. Ongelmaksi saattaa muodostua salaisten avainten turvallinen siirtäminen laitteesta toiseen käytettäessä eri laitteita. Toinen ongelma S/MIME-mekanismissa on, että se saattaa johtaa erittäin suurien viestien lähetykseen, varsinkin käytettäessä tunnelointia. Tunnelointia käytettäessä suositellaan käytettäväksi siirtoprotokollana TCP:tä [RFC 3261].

Loppukäyttäjien identifiointiin tarkoitetut S/MIME-sertifikaatit poikkeavat palvelinten käyttämistä sertifikaateista. Palvelinten sertifikaattien tarkistaessa tietyn toimialueen, loppukäyttäjien sertifikaatit varmistavat sertifikaatin

omistajan olevan identifioitu loppukäyttäjän osoitteella. Tämä osoite muodostetaan liittämällä ”käyttäjänimi”, ”'@” ja ”toimialue” osat SIP tai SIPS Internet-resurssin tunnisteesta. Nämä sertifikaatit liitetään avaimiin, joita käytetään salaamaan tai purkamaan SIP-viestien runkoja. Rungot ovat allekirjoitettu (signed) lähettäjän salaisella avaimella (private key) ja lisäksi ne ovat salattu vastaanottajan julkisella avaimella. Lähettäjä liittää viestiin oman julkisen avaimensa, jotta vastaanottaja pystyy lukemaan otsake-tietoja. Luonnollisesti lähettäjän tulee etukäteen tietää vastaanottajan julkinen avain, jotta salaus onnistuisi. Julkiset avaimet voidaan esimerkiksi tallentaa käyttäjäagenttiin [RFC 3261].

Käyttäjien tulisi hankkia sertifikaattinsa tunnetulta sertifikaatti-instanssilta. Vaihtoehtona sertifikaattien hankkimiselle on luoda sertifikaatti itse. Huolimatta sertifikaattien hankkimiseen liittyvistä ongelmista on olemassa muutamia tunnettuja keskushakemistoja, jotka jakavat sertifikaatteja loppukäyttäjille.

SIP:iäkin voidaan käyttää jakamaan julkisia avaimia. Käytettäessä “CMS SignedData” S/MIME -koodausta salaamaan SIP-viesti, sen tulee sisältää sertifikaatti. Tämän sertifikaatin tulee sisältää julkinen avain, jolla voidaan varmentaa allekirjoituksen aitous. Käytettäessä omatoimisesti luotuja sertifikaatteja tämä avainten vaihtomekanismi ei takaa luotettavaa avainten vaihtoa. Sen tarjoaman turvallisuuden katsotaan kuitenkin olevan ”parempi kuin ei mitään” [RFC 3261].

Seuraavassa koodiesimerkissä (koodiesimerkki 17) on esimerkki S/MIME-salatusta SDP-viestistä, joka on SIP-viestin rungossa.

```
INVITE sip:bob@biloxi.com SIP/2.0
Via: SIP/2.0/UDP pc33.atlanta.com;branch=z9hG4bKnashds8
To: Bob <sip:bob@biloxi.com>
From: Alice <sip:alice@atlanta.com>;tag=1928301774
Call-ID: a84b4c76e66710
CSeq: 314159 INVITE
Max-Forwards: 70
Contact: <sip:alice@pc33.atlanta.com>
Content-Type: application/pkcs7-mime; smime-type=enveloped-data;
        name=smime.p7m Content-Disposition: attachment;      Content-
Disposition: attachment; filename=smime.p7m handling=required

************************************************************
* Content-Type: application/sdp                            *
*                                                          *
* v=0                                                      *
* o=alice 53655765 2353687637 IN IP4 pc33.atlanta.com      *
* s=-                                                      *
* t=0 0                                                    *
* c=IN IP4 pc33.atlanta.com                                *
* m=audio 3456 RTP/AVP 0 1 3 99                            *
* a=rtpmap:0 PCMU/8000                                     *
************************************************************
```

Koodiesimerkki 17. S/MIME salattu viesti

Tarjotakseen jonkinlaista autentikointia, eheyttä ja luottamuksellisuutta SIP:in otsake-kentille alkupisteestä loppupisteeseen, S/MIME pystyy kapsuloimaan koko SIP-viestin ”message/sip”-tyyppisten MIME-runkojen sisään ja siten soveltamaan MIME:n tietoturvaa näihin runkoihin samaan tapaan kuin tavallisiin SIP-viestin runkoihinkin. Nämä kapsuloidut SIP-viestit eivät ole osana erillistä dialogia tai transaktiota, vaan ne ovat kopioita "ulkoisista" viesteistä, joita käytetään verifioimaan eheyttä tai tarjoamaan muuta lisätietoa [RFC 3261].

6.5. Mahdollisia tietoturvauhkia

SIP:in rekisteröintioperaation avulla käyttäjäagentit ilmoittavat mistä ne on tavoitettavissa. Rekisterinpitäjä tarkistaa ”From”- ja ”To”-kenttien perusteella, saako käyttäjä muokata haluttua sidontaa. Vaikka nämä kentät ovat usein samat, on olemassa keinoja, joita käyttäen kolmas osapuoli voi rekisteröityä

vastaanottamaan käyttäjän läsnäolotietoja. Jokainen voi itse muokata käyttäjäagenttinsa ”From”-kentän arvoa. Hyökkääjä, joka onnistuneesti tekeytyy osapuoleksi, jolla on oikeus muokata sidontoja, voi poistaa kaikki olemassa olevat sidonnat ja sen jälkeen lisätä oman päätelaitteensa ainoaksi sidonnaksi. Aiheuttaen näin kaikkien pyyntöjen toimittamisen hänen päätelaitteeseensa. Tämä uhkakuva on mahdollinen vain, mikäli viestien lähettäjien identiteettejä ei varmisteta. Yleensä kaikki palvelut, kuten rekisteröintipalvelu, haluavat kontrolloida palveluidensa käyttöä. Lisäksi useimmat palvelut ja päätelaitteet haluavat saada varmuuden toisen osapuolen identiteetistä [RFC 3261].

Toimialue, jossa SIP-pyynnön vastaanottaja sijaitsee, on yleensä määritelty SIP-viestissä. Käyttäjäagentit tai välipalvelimet ottavat yleensä suoraan yhteyden tämän toimialueen palvelimeen toimittaakseen pyynnön. Hyökkääjän on mahdollista yrittää tekeytyä täksi palvelimeksi ja näin saada haltuunsa pyyntö. Esimerkiksi uudelleenohjauspalvelin voi onnistuneesti tekeytyä toisen toimialueen uudelleenohjauspalvelimeksi, ja näin tämän toimialueen pyynnöt voidaan ohjata vääriin tai epäturvallisiin resursseihin. Pyynnöt voidaan myös jättää huomioimatta, aiheuttaen kaikkien pyyntöjen epäonnistuminen. Estääkseen tämän uhan toteutumisen käyttäjäagentin tulee autentikoida palvelin, jolle se lähettää pyyntöjä [RFC 3261].

SIP-käyttäjäagentit voivat käyttää viestien välittämiseen luotettuja välipalvelimia. Vaikka käyttäjäagentti luottaa välipalvelimen toimittavan viestit luotettavasti, se ei voi tarkistaa muuttiko tai lukiko välipalvelin viestin runko-osan. Esimerkki tälläisestä on käyttäjäagentti, joka käyttää SIP-viestien runkoja vaihtaakseen istuntokohtaisia salausavaimia. Se ei varmastikaan halua toimialueen ylläpitäjien mahdollisesti avaavan ja kuuntelevat istuntoa. Mikäli välipalvelin on vihamielinen, se voi muokata istunnon salausavainta toimien näin kahden osapuolen välissä (man-in-the-middle). Tämän kaltaiset uhat koskevat myös useimpia muita viestejä. Hyökkääjä voi esimerkiksi reitittää viestit kulkemaan palvelun kautta, joka tallentaa kaikkien viestien sisällön. Hyökkääjä voi muuttaa viestin otsikkokenttiä, kuten ”subject”-kenttää saaden viestin vaikuttamaan toiselta. Osan otsikkokentistä tulee aina olla selkokielisiä, koska niitä tarvitaan viestien toimittamiseen. Kaikkia otsikkokenttiä ei ole mahdollista turvata. Käyttäjäagentin on mahdollista turvata SIP-viestien runkoja sekä jossain määrin otsikkokenttiäkin koko viestin matkan ajaksi. Keinon tähän tarjoaa S/MIME [RFC 3261].

Istunnon perustamisen jälkeen sitä on mahdollista muokata. Istunnon osapuolten tulee olla varmoja istunnon muokkausten aitoudesta. Esimerkiksi kolmas osapuoli onnistuu kaappaamaan osan istunnon viesteistä saaden näin

haltuunsa istunnon tietoja (”To”-, ”From”- ja muita kenttiä). Tämän jälkeen hän lähettää BYE-viestin aiheuttaen istunnon ennenaikaisen lopettamisen. Tehokkain tapa torjua tämänkaltainen uhka on aina autentikoida BYE-viestin lähettäjä. Käyttäjän tulee tietää BYE-viestin tulevan istunnossa olevalta osapuolelta. Kolmas osapuoli ei voi väärentää BYE-viestiä, mikäli hän ei voi lukea istunnon tietoja [RFC 3261].

Palvelunestohyökkäys (denial-of-service) keskittyy estämään pääsyn tietylle verkkoelementeille. Yleensä tämä tapahtuu lähettämällä tietylle elementille suunnattomat määrät liikennettä ja jumiuttaen sen. SIP-välipalvelimet ovat yhteydessä Internetiin, jotta ne voivat toimittaa SIP-viestejä toisille toimialueille. Välipalvelin voi joutua palvelunestohyökkäyksen kohteeksi. Hyökkääjät voivat luoda väärät osoite- ja reititystiedot sisältäviä viestejä ja lähettää nämä viestit suurelle määrällä SIP-elementtejä. Paikkatietopalveluun voidaan luoda suuri määrä sidontoja, mikäli REGISTER-viestejä ei autentikoida ja aiheuttaa näin suuret määrät liikennettä. On myös mahdollista luoda riittävästi sidontoja, jotta paikkatietopalvelun fyysiset resurssit (muisti ja levytila) loppuvat [RFC 3261].

7. Yhteenveto

SIP on signalointiprotokolla ja sen tarkoitus on luoda, muokata ja lopettaa istuntoja. Se on suunniteltu toimimaan Internetissä ja yhdessä muiden web-teknologioiden kanssa. Vaikka SIP on melko uusi protokolla, sen pohjautuminen HTTP:hen luo sille lisätukea sen mekanismien toimivuudesta. SIP:iä kehitetään aktiivisesti IETF:n työryhmissä, mutta sen perusarkkitehtuuri on jo melko vakiintunut. SIP mahdollistaa laajennusten kehittämisen itse perus-SIP-arkkitehtuuriin, näin SIP:iä voidaan jatkuvasti kehittää lisäämällä siihen toiminnallisuutta.

Suoraviestintä on kahden tai useamman käyttäjän välistä viestintää lähes reaaliaikaisesti. SIP:iin on kehitetty suoraviestinnän mahdollistava laajennus, joka esittelee MESSAGE-viestin. SIP soveltuu hyvin suoraviestintään. MESSAGE-viesti mahdollistaa yksittäisten suoraviestien lähettämisen käyttäjältä toiselle. Se ei ota kantaa viestin sisältöön, joka voi olla mitä tahansa MIME-tyyppiä. Istuntopohjainen suoraviestintä käsitellään SIP:issä tavalliseksi istunnon luonniksi, johon SIP on alunperin tarkoitettukin.

Suoraviestintä SIP:issä ei kuitenkaan ole täysin ongelmatonta. Viestin lähettäminen vaatii aina SIP-otsakekentät, jolloin viestin koko kasvaa turhaan. Lisäksi koko SIP-arkkitehtuuri kohtaa ongelmia sen toimiessa verkossa, jossa on palomuureja sekä NAT-muunnoksia. Useimmat palomuurit eivät vielä tunnista SIP-liikennettä, jolloin ne eivät päästä sitä läpi. NAT:ien tehtävänä on muuntaa yritysten sisäisen verkon osoitteet toiseksi, joka lisää tietoturvaa. Näin SIP-viestintä ei välttämättä tiedä vastaanottajan oikeaa osoitetta.

Läsnäolotiedon liittäminen suoraviestintään voidaan ajatella teksti- ja mms-viestien seuraavaksi askeleeksi. Sen liittäminen olemassaoleviin sovelluksiin tarjoaa uusia mahdollisuuksia, joissa vastaanottaja on aina tavoitettavissa mikäli hän niin haluaa. Lisäksi se mahdollistaa paikkariippuvaisten palveluiden toteuttamisen. Läsnäolotiedon välittämiseen SIP soveltuu myöskin hyvin. Tosin ongelmaksi muodostuu SIP:in tekstipohjaisuus sekä viestien ”turhan” suuri koko. Tätä ongelmaa voidaan yrittää ratkaista pakkaamalla SIP-viestejä, jolloin lähetettävän datan määrä pienenee. Varsinaisia paikka-riippumattomia läsnäolotietojärjestelmiä ei ole vielä suuressa mittakaavassa toteutettu, joten niistä ei juuri ole kokemuksia.

Tietoturvaan SIP:issä on mielestäni kiinnitetty melko paljon huomiota. Tämä ennen kaikkea sen takia, koska SIP:in yleisin toimintaympäristö tulee olemaan Internet, jota pidetään vihamielisenä ympäristönä. Toisaalta voidaan miettiä, onko mikään Internetissä toimiva palvelu loppuen lopuksi riittävän

turvallinen. Tietoturva ei SIP:issäkään ole täysin ongelmatonta. SIP:in toiminnallisuuden takia, ei ole mahdollista salata kaikkia viestien otsakekenttiä, koska SIP-palvelinten tulee pystyä lukemaan ja muokkaamaan niitä. Muuten palvelimet eivät pysty toimittamaan viestejä vastaanottajille.

Kilpailu standardista suoraviestinnän sekä läsnäolotiedon esittämisestä tullaan epäilemättä käymään Jabberin (XMPP) ja SIP:in (SIMPLE) välillä. Molemmilla standardeilla on hyvät puolensa, joten on vaikea ilmaista kumpi on parempi. Dynamicsoftin päätutkija Jonathan Rosenberg, joka on ollut kehittä-mässä useita versioita SIP:istä, ilmaisi asian osuvasti:

> ”SIP is really good at rendezvous, call control, mobility support – none of which exists in Jabber / XMPP. Jabber is really good at carrying blocks of data between Point A and Point B, through firewalls, and with applications attached to the transport, like recording applications and so on. By putting them together to take advantage of Jabber for point-to-point messaging, it’s really the ideal combination.” [Rosenberg]

Jabber-yhteisö [Lear] näkee tulevaisuudessa neljä eri vaihtoehtoa SIP:in ja Jabberin välisessä suhteessa. Ne ovat:

- Jabber / XMPP katoaa markkinoilta. Tämä vaikuttaa erittäin epätodennäköiselta Jabberin suuren käytön takia.
- SIP / SIMPLE saa riittävästi markkinaosuutta, erityisesti Microsoftin ympäristöissä, joten se tulee elämään rinnakkain Jabberin kanssa. Yhteyskäytävät sekä palvelinpään lisäosat mahdollistavat niiden keskinäisen kommunikoinnin.
- Jabber / XMPP jatkaa kasvua ja SIP / SIMPLE ei saa markkinaosuutta, melkein syrjäyttää SIP / SIMPLE:n. SIP-laitteet alkavat kohtelemaan XMPP:tä kuten yhtä tiedonsiirtoprotokollaa ja käyttävät SIP:iä aloittamaan XMPP-pohjaisia istuntoja suoraviestinnän tarkoituksiin.
- Jabber syrjäyttää SIP:in, koska se pystyy keskustelemaan kaikkien XMPP-ratkaisujen kanssa. Tämä on mahdollista, muttei todennäköistä johtuen suuresta määrästä SIP-laitteita.

Itse uskon, että molemmat standardit tulevat jatkamaan eloaan ja kehittymään tulevaisuudessa. Uskon kuitenkin, että ne tulevat lähenemään toisiaan huomattavasti, ehkä jopa niiden välinen kommunikointi tulee ”suoraan” mahdolliseksi. Ainakin tullaan kehittämään välipalvelimia, jotka muuntaavat Jabber-viestit SIMPLE-viesteiksi ja päinvastoin. Näkemykseni on hieman samankaltainen kuin Jabber-yhteisön toinen vaihtoehto. Uskon kuitenkin, että SIP / SIMPLE tulee kasvamaan suuremmaksi ja käytetymmäksi

kuin Jabber. Perustan tämän uskomukseni sille, että Microsoft ja IBM ovat alkaneet tukea SIP:iä laitteissaan. Lisäksi 3GPP on valinnut SIP:in signalointiprotokollaksi 3G-verkoissa. Näiden seikkojen vuoksi SIP-pohjaisten ratkaisujen määrä tullee lähitulevaisuudessa kasvamaan. Lisäksi ohjelmistovalmistajien päätökset protokollan tukemisen aloittamisesta johtavat siihen, että se hyväksytään markkinoilla helpommin. Yritykset eivät halua käyttää ratkaisuja, joihin ne eivät saa tukea ongelmatilanteissa.

On kuitenkin huomioitava, että tällä hetkellä suoraviestintää hallitsevat suljetut standardit. Ei SIP eikä Jabber johda kilpailua avoimesta standardista suoraviestinnälle tai läsnäololle. Microsoftin päätös SIP:in käytön aloittamisesta johtaa siihen, että MSN Messenger alkaa jossain vaiheessa tukemaan SIP:iä. Jabber-yhteisö on myös kehittämässä yhdyskäytäviä, joiden avulla Jabber-suoraviestintäsovellukset voisivat keskustella muiden protokollien kanssa.

Viiteluettelo

[Camarillo, 2001] Camarillo, G., *SIP Demystified.* McGraw-Hill Professional Book Group, 2001.

[Cambpell et al.] Campbell, B., Rosenberg, J., Sparks, R., Kyzivat, P., The Message Session Relay Protocol, IETF Internet draft http://www.ietf.org/internet-drafts/draft-ietf-simple-message-sessions-03.txt. [23.3.2004]

[CPIM] Crocker, D., Diacakis, A., Mazzoldi, F., Huitema, C., Klyne, G., Rosenberg, J., Sparks, R., Sugano, H., Peterson, J., Common Profile for Instant Messaging (CPIM), IETF Internet draft, http://www.ietf.org/internet-drafts/draft-ietf-impp-im-04.txt. [23.3.2004]

[Koskela] Koskela, M., Internet Communications Using SIP: Presence and Instant communications. In: *Tietoliikenteen SIP seminaari.*

[Niemi] Niemi, A., Requirements for Instant Messaging in 3GPP Wireless Systems, IETF Internet draft, http://www.ietf.org/internet-drafts/draft-niemi-simple-im-wireless-reqs-02.txt. [23.3.2004]

[Lear] Lear, L., Comparing XMPP/Jabber and SIP/SIMPLE, Jabber White Paper,
http://www.jabber.com/media/1_The_IM_Standards_Race_v1.0.pdf
[23.3.2004].

[Liscano] Liscano, R., Presence and Awareness, http://www.site.uottawa.ca/~rliscano/tutorials/PresenceAwarenessServices.pdf.[23.3.2004]

[Osterman Research] http://www.ostermanresearch.com/research.htm [23.3.2004]

[RFC 768] Postel, J., User Datagram Protocol, RFC 768, http://www.ietf.org/rfc/rfc768.txt. [23.3.2004]

[RFC 791] Postel, J. (ed.), Internet Protocol - DARPA Internet Program Protocol Specification, RFC 791, http://www.ietf.org/rfc/rfc791.txt. [23.3.2004]

[RFC 793] Postel, J. (ed.), Transmission Control Protocol - DARPA Internet Program Protocol Specification, RFC 793, http://www.ietf.org/rfc/rfc793.txt. [23.3.2004]

[RFC 2616] Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Masinter, L., Leach, P. and Berners-Lee, T., Hypertext Transfer Protocol -- HTTP/1.1, RFC 2616, http://www.ietf.org/rfc/rfc2616.txt. [23.3.2004]

[RFC 2617] Franks, J., Hallam-Baker, P., Hostetler, J., Lawrence, S., Leach, P., Luotonen, A. and Stewart, L., HTTP Authentication: Basic and Digest Access Authentication, RFC 2617, http://www.ietf.org/rfc/rfc2617.txt. [23.3.2004]

[RFC 2633] Ramsdell. B. (ed.), H., S/MIME Version 3 Message Specification, RFC 2633, http://www.ietf.org/rfc/rfc2633.txt. [23.3.2004]

[RFC 2778] Day, M., Rosenberg, J., Sugano, H., A Model for Presence and Instant Messaging, RFC 2778, http://www.ietf.org/rfc/rfc2778.txt. [23.3.2004]

[RFC 2822] Resnick, P. (ed.), Internet Message Format, RFC 2822, http://www.ietf.org/rfc/rfc2822.txt. [23.3.2004]

[RFC 3261] Rosenberg, J., Schulzrinne, H., Camarillo, G., Jonston, A., Peterson, J., Sparks, R., Handley, M., Schooler, E., SIP: Session Initiation Protocol, RFC 3261, http://www.ietf.org/rfc/rfc3261.txt. [23.3.2004]

[RFC 3320] Price, R,. Bormann, C,. Christoffersson, J,. Hannu, H,. Liu, Z,. Rosenberg, J., Signalling Compression (SigComp), RFC 3320, http://www.ietf.org/rfc/rfc3320.txt. [23.3.2004]

[RFC 3428] Campbell, B., Rosenberg, J., Schulzrinne, H., Huitema, C., Gurle, D., Session Initiation Protocol (SIP) Extension for Instant Messaging, RFC 3428, http://www.ietf.org/rfc/rfc3428.txt. [23.3.2004]

[Rosenberg] Rosenberg, J., Presence and Awareness Services by Ramiro Liscano, http://www.site.uottawa.ca/~rliscano/tutorials/PresenceAwarenessServices.pdf page 64.[23.3.2004]

[Rosenberg 2] Rosenberg, J., A Presence Event Package for the Session Initiation Protocol (SIP), http://www.ietf.org/internet-drafts/draft-ietf-simple-presence-10.txt. [23.03.2004]

[Sinnreich and Johnston, 2001] Sinnreich H., Johnston, A.B., Internet Communications Using SIP: Delivering VoIP and Multimedia Services with Session Initiation Protocol. John Wiley& Sons, Inc, 2001.

Lyhenteet

Lyhenne	Merkitys	Selite
3G	3rd Generation	Kolmannen polven
AAA	Authentication, Authorization and Accounting	Autentikointi, todennus ja hallinnointi
B2BUA	Back-to-Back-User Agent	Erityinen käyttäjäagentti
CPIM	Common Profile for Instant Messaging	Viestiformaatti suoraviestinnälle
FTP	File Transfer Protocol	Tiedostonsiirtokäytäntö Internet-verkossa
HTTP	Hypertext Transfer Protocol	Verkkosivujen siirtokäytäntö
IETF	Internet Engineering Task Force	Internet-kehittämisryhmä
IMPP	Instant Messaging and Presence Protocol	Suoraviestinnän ja läsnäolon siirtokäytäntö
IMS	IP Multimedia Subsystem	Kolmannen polven televerkon osa
IPsec	IP Security	IP-verkon tietoturvastandardi
MGCP/Megaco	Media Gateway Control Protocol/Media Gateway Controller)	Puhelinverkon yhteyskäytäntö
MMUSIC	Multiparty Multimedia Session Control	SIP:in edeltäjä
MSRP	Message Session Relay Protocol	Istuntomoodin suoraviestin siirtokäytäntö
NAT	Network Address Translation	IP-osoitteiden muunto
PA	Presence Agent	Läsnäoloagentti
PIDF	Presence Information Data Format	Läsnäolotiedon viestiformaatti
PUA	Presence User Agent	Läsnäolotietokäyttäjäagentti
RFC	Request For Comments	Internetin käytäntöjen dokumentointi
RTP	Real-time Transport Protocol	Reaaliaikaisen audion ja videon siirtokäytäntö
SCIP	Simple Conference Invitation Protocol	SIP:in edeltäjä
SDP	Session Description Protocol	SIP istuntojen kuvauskieli
SIMPLE	SIP for Instant Messaging and Presence Leveraging Extensions	SIP laajennus suoraviestinnälle
SIP	Session Initiation Protocol	Signalointi siirtokäytäntö
SMTP	Simple Mail Transfer Protocol	Sähköpostin siirtokäytäntö
SS7	Signalling System 7	Yhteiskanavamerkinannon standardi
S/MIME	Secure/Multipurpose Internet Mail Extension	Turvallisempi MIME-sähköposti
UA	User Agent	Käyttäjäagentti
UAC	User Agent Client	Käyttäjäagentin asiakaselementti
UAS	User Agent Server	Käyttäjagentin palvelinelementti
UDP	User Datagram Protocol	Tiedonsiirtoprotokolla
URI	Uniform Resource Indicator	Internet-resurssin tunniste
VoIP	Voice over IP	Puheensiirto Internetissä
XML	eXtensible Markup Language	Merkkauskieli