



OULUN YLIOPISTO
UNIVERSITY of OULU

Sähköpostiprotokollat ylläpitävän organisaation näkökulmasta

Oulun yliopisto
Tietojenkäsittelytieteiden laitos
LuK-tutkielma
Saku Ruohola
18.3.2018

Tiivistelmä

Sähköpostijärjestelmät ovat tärkeä osa suurimman osan päivittäistä elämää. Erityisesti yrityselämässä sähköpostin rooli on edelleen merkittävä. Sähköpostiprotokollat ovat pysyneet suhteellisen muuttumattomina vuosikymmenten ajan, mutta ylläpitävien organisaatioiden työn määrä ei kuitenkaan ole samassa suhteessa vähentynyt. Nykypäivänä ongelmia aiheuttaa eritoten roskaposti, mutta myös kuluttajien muuttuvat tarpeet, joihin ratkaisuja haetaan käytettävistä ohjelmistoista protokollien sijaan.

Tarve tutkimukselle tuli ylläpitävältä organisaatiolta, joka haluaa päivittää käytössä olevaa sähköpostalustansa sen ollessa aikaansa jäljessä. Tästä syystä käsitellään ensin sähköpostijärjestelmiin liittyviä protokollia ja lopuksi vertaillaan tällä hetkellä käytössä olevia sähköpostalustoja parhaan vaihtoehdon löytämiseksi.

IMAP, POP3, SMTP, IMF, MIME, MUA, MSA, MTA, MDA

Valvoja

Ari Vesänen FT, yliopistonlehtori

Sisällysluettelo

Tiivistelmä	2
Sisällysluettelo	3
1. Johdanto.....	4
2. Tutkimusmenetelmä	5
3. IMAP- ja POP3-protokolla.....	6
3.1 IMAP	6
3.2 POP3	6
3.3 IMAP- sekä POP3-protokollan vertailu.....	7
4. Simple Mail Transfer Protocol (SMTP).....	9
5. Internet Mail Format (IMF).....	10
6. Multipurpose Internet Mail Extensions (MIME)	11
7. Sähköpostiviestinnän arkkitehtuuri	12
7.1 Mail User Agent.....	12
7.2 Mail Submission Agent	12
7.3 Mail Transfer Agent.....	14
7.4 Mail Delivery Agent	14
8. Sähköpostialustojen vertailu.....	15
8.1 Vertailukehikko	15
8.2 Vertailu eri alustojen välillä.....	16
9. Yhteenveto.....	18
Lähdeluettelo.....	19

1. Johdanto

Sähköpostiprotokollat eivät ole suuresti muuttuneet viimeisen vuosikymmenen aikana, mutta sähköposti on silti laajasti käytössä ympäri maailman ja erityisesti yrityksillä. Kuluttajatkin pitävät sähköpostia usein virallisempänä viestinnän välineenä kuin esimerkiksi sosiaalista mediaa. Sähköposti on yleisesti käytössä edelleen erittäin paljon myös markkinoinnissa. Vaikka sähköpostiprotokolliin ei ole tullut muutoksia aiheuttavat muuttuvat roskapostit sekä asiakkaiden vaihtelevat tarpeet silti palveluntarjoajalle alati haasteita.

Tässä tutkielmassa käydään läpi yleiset sähköpostiprotokollat sekä vertaillaan niiden toimintaa palveluntarjoajan näkökulmasta. Tarve tutkimukselle tuli palveluntarjoaja organisaation tarpeesta vaihtaa käytössä olevaa sähköpostalustaa. Tutkimusongelmana on siis selvittää sähköpostiprotokollien eroja palveluntarjoajan näkökulmasta. Tekninen tietoturva on liian laaja kokonaisuus käsiteltäväksi tässä tutkimuksessa, joten se rajataan tutkielmasta pois.

Tutkimusmenetelmänä käytetään kirjallisuuskatsausta. Käytetyt tietokannat, hakutermit sekä yhteenveto hakutuloksista on esitelty toisessa kappaleessa.

Kolmannessa kappaleessa tarkastelussa on Internet Message Access Protocol (IMAP) sekä Post Office Protocol version 3 (POP3). Erityisesti keskitytään niiden toiminnallisiin sekä hyötyihin ja vertaillaan niitä toisiinsa. Huomioitavaa on erityisesti erot palveluntarjoajan näkökulmasta. IMAP on protokollana uudempi kuin POP3, mutta molemmat ovat edelleen yleisesti käytössä, joten molempia käsitellään yhtä paljon.

Neljännessä kappaleessa esitellään yleisesti käytössä oleva sähköpostien välitysprotokolla Simple Mail Transfer Protocol eli SMTP. Käydään läpi toiminnallisuuksia joita protokollassa tulee olla ja joita erityisesti palveluntarjoajat tarvitsevat.

Viidennessä kappaleessa käsitellään Internet Mail Format, joka määrittelee tekstipohjaisten sähköpostiviestien syntaksin. Sen merkitys sähköpostin historiassa on merkittävä ja se on pysynyt suhteellisen muuttumattomana nykypäivään.

Kuudennessa kappaleessa esitellään Multipurpose Internet Mail Extensions, joka on edellisen kaltainen määrittely sähköpostiviesteille, mutta tekstipohjaisten viestien sijaan multimediatiedostoja sisältäville viesteille.

Seitsemännessä kappaleessa esitellään sähköpostijärjestelmässä käytettävät prosessit ja niiden vaikutus toisiinsa. Tarkoitus on käsitellä sähköpostijärjestelmää yleisestä näkökulmasta takertumatta teknisiin ratkaisuihin. Esitellään vaatimuksia joita protokollat aiheuttavat ylläpitäjälle sekä mahdollisia hyväksi todettuja käytäntöjä.

Kahdeksannessa kappaleessa esitellään ja vertaillaan eri sähköpostalustoja, joita tutkimuksen tilannut organisaatio voisi hyödyntää.

2. Tutkimusmenetelmä

Tutkimusongelmana on selvittää sähköpostiprotokollien merkitys ja erot ylläpitävän organisaation näkökulmasta. Tarve tutkimukselle tuli organisaatiolta, jonka on tarkoitus hyödyntää saatua tietoa sähköpostialustojen valintaa tehdessä.

Tutkimusmenetelmänä käytetään kirjallisuuskatsausta. Aineistoa on haettu seuraavista tietokannoista: Scopus, Google Scholar sekä IEEE/IEE Electronic Library. Suurin osa materiaalista löytyi Scopusen avulla.

Käytettyihin hakutermeihin sisältyi "IMAP", "Internet Message Access Protocol", "POP3", "Post Office Protocol", "SMTP", "Simple Mail Transfer Protocol", "email", "e-mail", "mail", "service provider", "compare", "comparison", "implement", "implementation", "MDA", "Mail Delivery Agent", "MTA", "Mail Transfer Agent", "MUA", "Mail User Agent", "MSA", "Mail Submission Agent", "MIME", "Multipurpose Internet Mail extensions", "IMF", "Internet Mail Format" sekä näiden yhdistelmiä. Haku rajattiin tietojenkäsittelytieteisiin, sillä tuloksia oli paljon. Hakusanat pohjautuivat alan terminologiaan ja ovat yleisesti käytössä.

Tarkastellaan sähköpostiprotokollia erityisesti ylläpitävän organisaation näkökulmasta. Siksi hakulausekkeissa on huomioitu palveluntarjoaja, mutta tarpeen on myös taustatiedon kartuttaminen eri protokollien osalta.

Tämän lisäksi aineistona on käytetty RFC eli "Request for comments"-dokumentteja jotka toimivat eräänlaisina määrittelydokumentteina protokollille. RFC -dokumentteihin on viittauksia useissa käytetyissä lähteissä. RFC -dokumentit ovat julkista tietoa ja löytyvät Google -hakukoneen avulla.

3. IMAP- ja POP3-protokolla

Internet Message Access Protocol sekä Post Office Protocol version 3 ovat tärkeimmät ja käytetyimmät protokollat sähköpostien hakemiseen palvelimelta.

3.1 IMAP

Internet Message Access Protocol eli lyhennettynä IMAP on sähköpostiprotokolla, joka mahdollistaa sähköpostien lukemisen, lähettämisen sekä hallinnoinnin, mutta ei toimi sähköpostien välittäjänä. Sähköpostien välittämiseen on omat protokollansa joista esimerkkinä kolmannessa kappaleessa läpikäytävä SMTP (RFC3501, 2003).

IMAP-protokollan avulla voi sähköpostilaatikossa luoda sekä poistaa viestejä, uudelleen nimetä sähköpostilaatikoita, hakea uusia viestejä, poistaa pysyvästi viestejä, hakea saapuneista viesteistä hakusanalla, suodattaa tulevia viestejä sekä asettaa tai poistaa jo asetettuja lisätietoja viesteistä (RFC3501, 2003). IMAP mahdollistaa saman sähköpostilaatikon käyttämisen useammalla eri laitteella, sillä sähköpostit tallentuvat palvelimelle ja niistä haetaan kopiot käytettävälle laitteelle (Kissell, 2006). Tehdyt muutokset näkyvät kaikille käyttäjille ja niitä voi tehdä, vaikkei laite olisi yhteydessä palvelimeen. Muutokset astuvat voimaan ja siten näkyvät kaikille käyttäjille kuitenkin vasta yhteyden muodostuttua (RFC3501, 2003).

Viestin luomisella tarkoitetaan uuden sähköpostiviestin kirjoittamista. IMAP-protokollan avulla viesti voidaan poistaa paikallisesti sitä käytettävästä sovelluksesta tai pysyvästi jolloin se poistuu myös palvelimelta. Saapuvien viestien suodattaminen onnistuu myös IMAP-protokollasta löytyvillä toiminnallisuuksilla sekä jo saapuneista viesteistä hakeminen erilaisilla hakuehdoilla esimerkiksi lähettäjän tiedoilla. Viestien tunnisteita on mahdollista asettaa sekä poistaa. Tunnisteita ovat esimerkiksi tieto vasta saapuneesta sähköpostista sekä siitä onko se jo luettu. Näitä kaikkia on siis mahdollista muokata eli esimerkiksi jo luetun viestin voi merkata lukemattomaksi. (RFC3501, 2003).

IMAP protokolla on ylläpitävälle organisaatiolla raskas, sillä sähköpostit säilytetään pääosin palvelimella ja käyttäjät luottavat sen olevan kaiken aikaa käytettävissä. Käyttäjät voivat halutessaan ladata sähköpostit omille laitteilleen ja vapauttaa siten tilaa palvelimelta. Sähköpostien säilyttäminen palvelimella hyödyttää erityisesti käyttäjiä joilla on hidas internetyhteys, sillä suurten sähköpostien ja niiden liitetiedostojen lataaminen omalle laitteelle kuormittaa internetyhteyttä. Palvelimen virhetilanteissa käyttäjät eivät kuitenkaan pääse käsiksi sähköposteihin, joita eivät ole ladanneet omille laitteilleen, mikä aiheuttaa ylläpitäjille lisävastuuta (Dean, 2009).

3.2 POP3

Post Office Protocol version 3 eli lyhennettynä POP3 on yksinkertainen sähköpostiprotokolla, joka mahdollistaa sähköpostien lukemisen, lähettämisen sekä poistamisen. POP3-protokolla on tarkoitettu käytettäväksi yhdellä laitteella, sillä oletuksena sähköpostit ladataan käytettävälle laitteelle eikä palvelimelle jää kopioita. POP3 mahdollistaa kuitenkin sähköpostien säilyttämisen palvelimella, jolloin samaa

laatikkoa voi käyttää useammalla laitteella, muttei silti samanaikaisesti (RFC1939, 1996). Sähköpostien välittämiseen tulee käyttää erillistä protokollaa, joista esimerkkinä SMTP on esitelty kappaleessa kolme.

POP3 on tarkoituksella yksinkertaistettu sähköpostiprotokolla ja sen määrittelyssä onkin mainittu, että mikäli tarvetta on monimutkaisemmalle protokollalle tulisi käyttää IMAP-protokollaa eikä tehdä muutoksia POP3-protokollaan (RFC2449, 1998). POP3-protokollaa tulisikin käyttää vain, mikäli pelkkä viestien hakeminen ja poistaminen riittävät (RFC1939, 1996).

POP3-protokolla sallii palveluntarjoajan asettaa paikallisia sääntöjä esimerkiksi viestien poistamiseen. Palveluntarjoaja voi esimerkiksi halutessaan asettaa kaikki yli 60 päivää vanhat viestit poistumaan. Palveluntarjoajan tulee kuitenkin tiedottaa tästä käyttäjille erityisen hyvin, sillä käyttäjän käyttämä sovellus voi antaa ymmärtää viestien säilyvän palvelimella ikuisesti (RFC1939, 1996).

POP3-protokolla käyttää selvästi vähemmän palvelimen resursseja kuin IMAP-protokolla, joten se on palveluntarjoajan näkökulmasta parempi. POP3-protokollalle on myös tuki lähes kaikilla laitteilla (Dean, 2009)

3.3 IMAP- sekä POP3-protokollan vertailu

Taulukko 1. Vertailu IMAP- sekä POP3-protokollan ominaisuuksista.

	IMAP	POP3
Nykyinen versio	IMAP4rev1	POP3
Sisällön tallennus	Aina palvelimella	Paikallisella laitteella, ellei erikseen valita palvelinta
Useampi samanaikainen käyttäjä	Kyllä	Ei
Salaus	Mahdollinen	Mahdollinen
Palvelimen käyttämä portti	143 ja salattuna 993	110 ja salattuna 995
Kansiointi	Kyllä	Ei
Synkronointi	Kyllä	Ei
Hakujen tekeminen palvelimelta	Kyllä	Ei
Julkaisuvuosi	1990	1988

Sisällön tallennus eli sähköpostien sekä niiden liitteiden säilytys on IMAP-protokollassa aina palvelimella, josta sitten haetaan kopiot käytettävään laitteeseen. POP3-protokollassa oletuksena tiedot haetaan aina käytettävään laitteeseen, mutta ne on

mahdollista säilyttää myös palvelimella, jolloin voidaan käyttää useampaa laitetta eri aikoina. (RFC3501, 2003; RFC1939, 1996).

IMAP sallii useamman käyttäjän tai laitteen käyttää sähköpostilaatikkoa samanaikaisesti. POP3 -protokollassa ei ole tuettu samanaikaista käyttöä eikä siinä ole sitä tukevia toimintoja (RFC3501, 2003; RFC1939, 1996).

Molemmilla protokollilla on oletuksena omat tietoliikenneportit, mutta IMAP-protokolla voi käyttää myös muita portteja (RFC3501, 2003; RFC1939, 1996).

IMAP-protokolla sallii useampien sähköpostilaatikoiden hallinnan ja niiden välillä sähköpostien vaihtamisen (RFC3501, 2003). Käyttäjä voi jakaa sähköpostilaatikoitaan julkisina kansioina ja rajoittaa niiden näkyvyyttä (RFC4314, 2005). POP3-protokollassa ei ole mahdollista hallita useampaa sähköpostilaatikkoa.

Sähköpostiviestien tilatietoja ei ole POP3-protokollassa ja siten esimerkiksi tietoa onko viesti luettu ei voida synkronoida eri laitteiden välillä eri aikoihin kuten IMAP-protokollassa (RFC3501, 2003). POP3 ei myöskään tue uusien viestien synkronointia palvelimelta kuten IMAP ja siten käyttäjä ei tiedä uusista viesteistä, ellei erikseen avaa POP3 sähköpostilaatikkoaan (RFC1939, 1996).

IMAP-protokollassa voidaan hakuja tehdä palvelimelta, jolloin kaikkia viestejä ei tarvitse ladata käytettävään laitteeseen. Tämä ei ole mahdollista POP3-protokollassa. Hakujen tekeminen on kuitenkin riippuen palvelimen hakuoptimoinnista usein erittäin raskasta palvelimelle varsinkin isojen sähköpostilaatikoiden tapauksessa (RFC3501, 2003).

Tehokkuus protokollien välillä vaihtelee näkökulmasta riippuen. POP3-protokollassa kaikki sähköpostit ladataan oletuksena käytettävälle laitteelle ja siten sen kuormitus palvelimelle ja sitä kautta ylläpitävälle organisaatiolle on huomattavasti pienempi kuin IMAP-protokollan. Sen sijaan kuormitus käyttäjän laitteelle sekä Internet-yhteydelle on POP3-protokollan tapauksessa suurempi. Palveluntarjoajan näkökulmasta IMAP-protokolla on siis selvästi raskaampi kuin POP3, sillä sen tarjoamat ominaisuudet käyttävät selkeästi enemmän resursseja palvelimelta (RFC3501, 2003; RFC2449, 1998).

Sähköpostien tallennus on palveluntarjoajan näkökulmasta suurimpia resurssien käyttäjiä ja siten sen tehokkuuteen panostaminen on tärkeää ylläpitäjälle. Raghunandan esittää tutkimuksessaan uudenlaista lähestymistä kapasiteettiongelmaan. Hänen ratkaisussaan palveluntarjoajat voisivat lainata toisilleen tallennustilaa tarvittaessa. Tämän lisäksi kuluttajan sähköposteja voitaisiin pakata tilan säästämiseksi tietyn rajan ylityessä. Sähköpostit purettaisiin vain kun käyttäjä haluaa avata niitä (Raghunandan, 2009).

Vuonna 2011 Srinivasan tekemässä tutkimuksessa testattiin sähköpostipalveluiden käyttämistä henkilökohtaisena pilvipalveluna tiedostojen tallennukseen. Sähköpostilaatikoiden tilat ovat kasvaneet siinä missä suurin osa pilvipalveluista on maksullisia tai rajoittaa käyttäjien käyttämiä sovelluksia. Tutkimuksessa huomattiin etteivät sähköpostilaatikat ole kuluttajille saatavilla olevia pilvipalveluita huonompia. Tämä trendi on kuitenkin yleistyessään huolestuttava sähköpostipalveluntarjoajan näkökulmasta, sillä sähköpostille varattuja palvelimia ei ole tarkoitettu tiedostojen säilyttämiseen (Srinivasan, 2011).

4. Simple Mail Transfer Protocol (SMTP)

SMTP eli Simple Mail Transfer Protocol on luotettava ja tehokas sähköpostien välitysohjelma (RFC5321, 2008). Ohjelmalla on alustariippumaton ja siten se vaatii ainoastaan siirtotien tiedolle (Karvinen, 1997). SMTP on tarkoituksella yksinkertainen ohjelma, jonka ainoa tehtävä on siirtää sähköposteja sekä säilyttää niitä jonossa odottamassa siirtoa (Dean, 2009).

Sähköpostijärjestelmän luotettavuuden sekä ennustettavuuden takia kaikkien lähetettyjen sähköpostien tulisi saapua vastaanottajalle tai mikäli se ei ole mahdollista pitää siitä tulla ilmoitus lähettäjälle. Suurin osa sähköpostiliikenteestä on kuitenkin roskapostia ja siten tieto lähettäjälle saattaa olla haitallista lisääntyvän roskapostin takia. Ilmoittamaton sähköpostien poistaminen vaarantaa sähköpostien luotettavuutta ja siksi sitä tulisi käyttää vain selkeimpiin roskaposteihin. Sähköpostiviestien sisältöön ja niiden haitallisuuteen SMTP –ohjelmalla ei ota kantaa eikä sen määrittelyssä ole siis huomioitu sähköpostien sisältöä millään tavalla (RFC5321, 2008). Tämä tarkoittaa sitä ettei SMTP -ohjelmalla nykyisellään pyri korjaamaan virheellisiä viestejä vaan yrittää yksinkertaisesti lähetystä uudelleen virheen sattuessa (Dean, 2009).

SMTP –ohjelmalla vaatii palvelimelta silmukoiden esto eli automaattisten viestien loppumattoman ristiin lähettämisen keskeyttämistä. Yksinkertaisimmillaan tämä voidaan toteuttaa laskemalla ”saapunut”-kentän arvoa. Tämä ei toki ole kovinkaan tehokas keino ja kentän arvo tuleekin määrittää riittävän korkeaksi, ettei oikeita viestejä estettä (RFC5321, 2008).

SMTP-ohjelmalla perustamisesta lähtien on käyty keskustelua tulisiko ohjelmalla pyrkiä virheellisten viestien korjaamiseen vai suoraan hylätä virheellisiä viestejä. Viestien korjaamista puoltavat perustelevat tätä käyttäjien toiveena saada viestit lähetettyä, jos se on mitenkään mahdollista. Viestien hylkäämistä perustellaan tehokkaana sekä ainoana keinona saada virheelliset järjestelmät sekä sovellukset korjattua (RFC5321, 2008). Virheiden korjaamiseen ja käsittelyyn voidaan käyttää POP3 tai IMAP -ohjelmia, jotka on esitelty aikaisemmassa kappaleessa (Dean, 2009).

SMTP –ohjelmalla tulee olla toteutus ajastetulle automaattiselle viestin uudelleenlähetykselle, mikäli viestin lähetyksessä jostain syystä epäonnistuu. Tehokkaana pidetään ratkaisua, jossa uudelleenlähetyksen viive vaihtelee saadun virheilmoituksen mukaan. Uudelleenlähetyksiä ei saa tehdä loputtomiin, mutta hyvänä käytäntönä on pidetty neljästä viiteen päivää, jonka jälkeen lähettäjälle ilmoitetaan lähetyksen epäonnistumisesta. Kaikkien uudelleenlähetyksiin liittyvien aikojen tulee olla muokattavissa (RFC5321, 2008).

SMTP –ohjelmalla tulisi kaiken aikaa ylläpitää yhteyttä porttiin 25 sähköpostiviestien vastaanottamisen vuoksi. Tämä vaatii useampaa samanaikaista yhteyttä, mutta niiden rajoittaminen on ohjelmalla sallittua. Palvelin joka ei pysty kuin yhteen samanaikaiseen yhteyteen ei ole SMTP –ohjelmalla mukainen (RFC5321, 2008). Vuonna 1997 julkaistussa tutkimuksessa Tzerefos ja kumppanit todistivat, että yhteyden ylläpitäminen kaikkien viestien välittämiseen asti lisää merkittävästi ohjelmalla tehokkuutta ja jo siitä syystä ohjelmalla käyttävän tulisi mahdollistaa riittävien yhteyksien luominen (Tzerefos et al, 2009).

5. Internet Mail Format (IMF)

Internet Mail Format (IMF) on RFC822 -dokumentissa ensimmäisen kerran määritelty syntaksi tekstipohjaisille sähköposteille. Sitä on sittemmin päivitetty aina nykyiseen RFC5322 versioon sekä siihen tehtyyn lisäykseen RFC6854. IMF ei ota kantaa kuvien, äänitiedostojen tai muiden tiedostotyyppien lähetykseen, mutta näistä on tehty erillisiä dokumentteja, joista yksi esitellään seuraavassa kappaleessa (RFC5322, 2008).

Sähköpostiviesti koostuu eräänlaisesta kirjekuoresta sekä sen sisällöstä. Kirjekuorella tarkoitetaan lähettäjän sekä vastaanottajan tietoja sekä mahdollisia muita kenttiä, joita tarvitaan viestin välittämiseen SMTP -protokollalla. Viestin sisältö koostuu otsikosta sekä rungosta (RFC5322, 2008).

Viestin otsikon pituuden ei suositella olevan yli 78 merkkiä eikä se saa olla yli 998 merkkiä pitkä. Sama rajoitus koskee viestin runkoa siten ettei yksittäinen rivi rungossa saa olla yli 78 merkkiä eikä sen suositella olevan yli 998 merkkiä pitkä. Merkistönä käytetään US-ASCII:ta (RFC5322, 2008).

Syntaksi sallii erilaisten otsikkotietojen käyttämisen, kuten esimerkiksi päivämäärän sekä tarkan ajan lisäämisen sähköpostiin. Ajan lisäksi välitetään tieto käytettävästä aikavyöhykkeestä ja siksi suositellaankin käytettävän paikallista aikaa. Erilaisten otsikkotietojen käyttämiseen on tarkat määritelmät ja ne tulee olla oikeassa muodossa, jotta tieto välittyy oikein eri sähköpostipalveluiden välillä. Pakollisia kenttiä ovat vain lähettäjä ”from” sekä lähetys päivämäärä ”orig-date”, mutta virheiden välttämiseksi suositellaan myös lähettäjäkenttää ”sender”, viestin identifikaatiokenttää ”message-id”, viestiketjun identifiointikenttää ”in-reply-to”, uudelleenlähettäjäkenttää ”resent-sender” sekä viittauskenttää ”references” mikäli viestissä on sellaista mahdollista käyttää (RFC5322, 2008).

Otsikkokenttänä voi siis olla myös tunnistuskenttiä, joiden avulla vastaukset voidaan liittää oikeaan sähköpostiketjuun. Tämän lisäksi voidaan esimerkiksi käyttää erillistä ”sender” -kenttää ”from” -kentän lisäksi kun tarve on lähettää sähköposteja jonkun nimissä (RFC5322, 2008).

6. Multipurpose Internet Mail Extensions (MIME)

Multipurpose Internet Mail Extensions (MIME) on edellisessä kappaleessa käsitellyn IMF -syntaksin kaltainen määrittely sähköpostiviestien sisällölle. Sen tarkoitus on lisätä tuki myös ASCII merkistön ulkopuolisille merkistöille sekä mahdollistaa muiden kuin tekstien käyttäminen viestin rungossa. Se mahdollistaa myös useampi osaisen viestirungon, joka mahdollistaa pitempien sähköpostien lähettämisen (RFC2046, 1996).

Suurimpia rajoitteita IMF -syntaksissa on tuki pelkästään US-ASCII merkistölle sekä viestien pituuden rajoittaminen alle tuhanteen merkkiin. Tämän lisäksi IMF ei ota millään tavalla kantaa muihin kuin tekstipohjaisiin viesteihin (RFC2045, 1996).

MIME -syntaksi lisää sähköpostiin uusia otsikkokenttiä, joiden avulla erilaisia tiedostotyyppjä voidaan tunnistaa. Myös versiolle on oma kenttänsä, jotta ohjelmat tunnistavat sähköpostin toteuttavan MIME -syntaksia. Nämä uudet otsikkokentät ovat IMF -syntaksin mukaisia, mutta niitä käytetään siis vain MIME:n kanssa (RFC2045, 1996).

MIME:n tarkoitus on välttää yhteensopivuusongelmia ja siten erilaiset versiot ja niiden hallinta on otettu huomioon määrittelyssä. Se toimii pitkälti samalla tavalla kuin IMF, mutta lisää uusia toiminnallisuuksia tiedostotyyppien sekä merkistöjen osalta (RFC2045, 1996).

MIME lisää tuen kuva-, ääni-, video- sekä sovellustiedostoille sähköpostiviesteissä. Viesteihin on myös mahdollista määritellä useampi osaisia tiedostoja (RFC2046, 1996).

7. Sähköpostiviestinnän arkkitehtuuri

Tässä luvussa tarkastellaan tarkemmin sähköpostiviestinnässä käytettäviä prosesseja sekä niiden vaikutusta toisiinsa. Kappaleen tarkoitus on syventää ymmärrystä erityisesti eri prosessien vastuista, jotta seuraavassa kappaleessa eri ohjelmien vertailuun on riittävästi tietoa.

7.1 Mail User Agent

Mail User Agent on sähköpostien kohde sekä lähettäjä. Sillä tarkoitetaan siis käyttäjää tai usein sen käyttämää sovellusta. Mail User Agent huolehtii siis sähköpostien siirtymisestä Message Transfer Agent:ille (MTA) ja toki myös niiden vastaanottamisesta palvelimelta eli Mail Delivery Agent:lta (MDA) SMTP:n avulla. Raja MUA:n sekä MTA:n välillä ei aina ole kovinkaan selkeä, sillä erinäiset sovellukset voivat saada molempia rooleja. Tämän lisäksi raja vastuissa ei ole täysin yksiselitteinen, mikä osaltaan hankaloittaa linjavetojen tekemistä (RFC5321, 2008).

Käyttäjällä voidaan eri tilanteissa tarkoittaa ihmisten lisäksi esimerkiksi organisaatioita sekä erilaisia prosesseja. Käyttäjät jaetaan neljään pääryhmään jotka ovat (RFC5598, 2009):

1. Kirjoittaja, joka luo viestin sekä valitsee sille vastaanottajat.
2. Vastaanottaja, joka vastaanottaa kirjoittajan tekemän viestin sekä voi halutessaan vastata siihen.
3. Palautusten käsittelijä, joka käsittelee mahdollisten virheiden aiheuttamat vastaukset oikeille vastaanottajille eli usein lähettäjälle.
4. Välittäjä, joka vastaanottaa ja mahdollisesti muokkaa viestiä ennen uudelleen lähetystä. Usein viestistä pyritään säilyttämään mahdollisimman paljon ja muutoksia tehdään siis mahdollisimman vähän. Näistä esimerkkeinä ovat esimerkiksi postituslistat, jotka toimivat välittäjinä useisiin osoitteisiin. Myös monimutkaisemmat prosessit ovat mahdollisia.

Mail User Agent voi myös käyttää erillistä pitkäaikaisempaa varastoa viesteille, jota kutsutaan yksinkertaisesti nimellä Message Store (MS). MS voi sijaita erillisellä palvelimella johon MUA ottaa erikseen yhteyttä, mutta se voi myös olla käyttäjän koneella. MUA voi hakea viestejä lähettäjältä proaktiivisesti erillisellä toiminnolla tai SMTP-, POP3- tai IMAP-protokollan avulla. MUA ottaa yhteyden viestivarastoon joko erillisellä toiminnolla tai POP3- tai IMAP -protokollalla. Viestien yksittäinen siirtäminen POP3 -protokollalla ei kuitenkaan ole kovinkaan tehokasta ja siitä syystä se ei olekaan laajasti käytössä (RFC5598, 2009).

7.2 Mail Submission Agent

Mail Submission Agent:a (MSA) voidaan käyttää lisäämään turvallisuutta sähköpostien käsittelyssä. Mikäli sitä käytetään lähettää MUA viestit MTA:n sijaan MSA:lle, joka

puolestaan käsittelee ne mahdollisesti suodattaen sekä lähettää ne sitten eteenpäin MTA:lle tai suoraan vastaanottajalle (RFC6409, 2011). Mail Submission Agent käyttää porttia 587, mutta sen voi tarvittaessa vaihtaa käyttämään porttia 25. MTA:lle sekä MSA:lle on mahdollista asettaa erilaiset suodattimet, jolloin voidaan paremmin hyödyntää esimerkiksi paikallista autentikointia (RFC6409, 2011). Käyttäjien autentikointiin on useita mahdollisia keinoja, kuten esimerkiksi autentikoitu SMTP sekä turvattu IP-osoite. Autentikoitu SMTP on yleisesti käytössä ja mahdollistaa MSA:n varmistaa lähettäjä ilman erillistä protokollaa. Erilaiset rajoitteet IP-osoitteisiin ovat myös suuresti käytössä, mutta IP-osoitteiden helpon väärentämisen takia tätä ei pidetä erittäin luotettavana. IP-osoitepohjaiset rajoitukset toimivatkin vain mikäli kaikki välitykseen käytettävät palvelimet ovat luotettavia. (RFC6409, 2011).

Jokaisen MSA:n tulee tehdä kaikki seuraavista (RFC6409, 2011):

1. Lähettää hylkäysviesti, mikäli sellaiselle tulee tarvetta. Oletuksena tämän virhekoodi on 554, mutta tarvittaessa voidaan käyttää myös tarkennettuja hylkäysviestejä eri koodeilla.
2. Varmistaa että kaikki käytettävät verkkotunnukset ovat esteettömiä eli niihin on pääsy. Mikäli näin ei ole tulee tästä lähettää hylkäysviesti yllä mainitulla tavalla.
3. Vaatia autentikointi. Oletuksena MSA:n tulee vaatia autentikointi ja lähettää lähettäjäälle virhe koodilla 530, mikäli autentikointi puuttuu.

Lisäksi suositellaan seuraavia toimintoja (RFC6409, 2011):

1. Hylätä viestit joiden lähettäjä tai vastaanottaja -kenttien tiedot eivät ole oikein. Tähän suositellaan virhekoodia 501.
2. Kirjata mahdolliset virhetilanteet erityisesti liittyen väärin konfiguroituihin asiakasohjelmistoihin. Tästä on erityisesti hyötyä ylläpitävälle taholle, sillä he voivat hoksauttaa käyttäjiä tai heidän ylläpitäjiään.
3. Säättää aikakatkaisuihin vaadittavia aikoja lyhyemmiksi, sillä palvelinten tulisi todellisuudessa vastata mihin tahansa komentoon alle parissa minuutissa.

MSA saa myös tehdä mitä tahansa seuraavista (RFC6409, 2011):

1. Eritellä oikeuksia ja rajoittaa käyttöä niiden avulla.
2. Valvoa asettamia rajoituksia ja lähettää tarvittaessa hylkäysviesti esimerkiksi koodilla 550.
3. Tarkistaa viestin sisältöä syntaksin ja mahdollisten ristiriitojen suhteen. Ristiriitoja voi tulla esimerkiksi käyttäjän käyttöoikeuksien kanssa. Koodia 550 suositellaan käytettävän jos hylkäys liittyy käyttäjään, kun taas koodia 554 jos vika on viestin syntaksissa ja koodia 501 jos komento itsessään ei ole syntaksin mukainen.
4. Tukea ”Postmaster” -osoitetta eli ylläpitäjälle tarkoitettua osoitetta, jota voidaan käyttää virheiden raportoimiseen.
5. Muuttaa käytettäviä merkistöjä, jotta viestit menisivät varmemmin perille.

7.3 Mail Transfer Agent

Mail Transfer Agent toimii SMTP-palvelimena, joka saa viestejä MSA:lta tai toiselta MTA:lta ja välittää niitä toisille MTA palvelimille tai kuljettaa ne suoraan vastaanottajalle (RFC6409, 2011). Roskaposti on yleistynyt sähköpostin alkuajoista asti ja nykyään valtaosa sähköpostiliikenteestä koostuu roskapostista. Tästä syystä on tehty erikseen dokumentteja roskapostin suodattamiseen nimenomaan MTA:lle tehtävien muutosten avulla. Esimerkiksi RFC 2505 käsittelee pelkästään hyväksi todettuja käytäntöjä roskapostien estämiseen. Roskaposti aiheuttaa merkittävää haittaa ylläpitävälle organisaatiolle, sekä palveluntarjoajalle, koska se täyttää sähköpostipalvelimilla sähköposteille varatut tilat nopeasti ja estää näin käyttöä. Tästä syystä roskaposti tulee ottaa ehdottoman vakavasti (RFC2505, 1999).

7.4 Mail Delivery Agent

Mail Delivery Agent siirtää vastuun viestien käsittelijältä vastaanottajalle eli useimmissa tapauksissa kuljettaa viestin lähettäjältä vastaanottajan palvelimelle ja tämän sähköpostilaatikkoon. MDA:n tulee tietää vastaanottaja ja sillä tulee olla jokin keino saavuttaa vastaanottaja. Siirto MDA:lle tapahtuu ylläolevan Message Transfer Agentin avulla ja edelleen siltä käyttäjälle joko POP3 tai IMAP protokollalla (RFC5598, 2009). Sekä MDA:n että MSA:n on mahdollista estää sähköpostien välittäminen mikäli lähettäjä ei ole ennestään tuttu. Tätä pidetäänkin hyvänä käytäntönä, sillä välittäjien olisi hyvä pitää kirjaa ja tarvittaessa rajoittaa käyttäjien lähetyksiä (RFC5068, 2007).

8. Sähköpostialustojen vertailu

Tutkimusongelmana oli löytää uusi vaihtoehtoinen sähköpostialusta tutkimuksen tilanneelle organisaatiolle. Erilaisia sähköpostialustoja löytyi Google -hakukoneen avulla useita. Löytyneistä sähköpostialustoista karsittiin kuluttajille suoraan tarkoitetut alustat ja keskityttiin erillisten organisaatioiden ylläpidettäviin alustoihin. Alla on vertailukehikko eri ominaisuuksista joita mielestäni alustassa tulisi olla. Tämän jälkeen on vertailtu löydettyjä sähköpostialustoja kuvaillun vertailukehikon avulla.

8.1 Vertailukehikko

Tutkimuksen tilanneella organisaatiolla tärkein ominaisuus on sähköpostialustalle on pieni kustannus, mutta muut ominaisuudet vertailukehikkoon on kirjoittaja itse päättänyt ja siten ne saattavat poiketa eri organisaatioiden tarpeista. Ominaisuuksia vertailukehikkoon on saatu ylläpitävältä organisaatiolta. Tämän lisäksi on valittu muutamia mielestäni tärkeitä ominaisuuksia.

Ilmainen: Ilmaisuus on ylläpitävän organisaation näkökulmasta erittäin tärkeää, sillä asiakkaita voi olla useita ja rahoitusmallit ovat näissä tapauksissa usein hankalia. Näistä esimerkkinä Zimbra, jonka rahoitusmalli on sidottu käytettyjen sähköpostilaatikkojen määrään. Rainloop:illa oli sekä ilmainen että maksullinen versio, joista taulukossa on käytetty ilmaista Community -versiota.

Selainpohjainen: Selainpohjaisuus on yksi tutkimuksen tilanteen organisaation kriteereistä ja sillä se otettiin taulukossa vertailuksi. Selainpohjaisuudella tarkoitetaan yksinkertaisesti sitä ettei sähköpostin käyttämiseksi tarvitse ladata erillisiä ohjelmia, vaan käyttö onnistuu jollain yleisistä selaimista.

Mobiilituki: Mobiilituella tarkoitetaan selainnäköymän mukautumista myös pieniresoluutioisille laitteille kuten puhelimille sekä tableteille. Tämä on monille käyttäjille tärkeä ominaisuus ja siten myös palveluntarjoajien näkökulmasta tärkeä.

POP3 sekä IMAP -tuki: POP3 sekä IMAP -tuella viitataan ohjelman sallimiin protokollisiin. Eli onko käyttäjän mahdollista luoda kyseistä protokollaa käyttävä sähköpostilaatikko. Palveluntarjoajan näkökulmasta tuki molemmille olisi hyvä, vaikkei käyttäjät välttämättä osaisi molempia kaivata.

Sähköpostien suodatus: Sähköpostien suodattaminen tarkoittaa käyttäjän asettamia ehtoja joiden mukaan saapuvia sähköposteja ohjataan käyttäjän määrittelemien asetusten mukaan esimerkiksi tiettyihin kansioihin.

Sähköposteista hakeminen: Sähköposteista hakeminen tarkoittaa saapuneista viesteistä yksittäisten tai useiden viestien hakemista tietyillä käyttäjän valitsemilla hakusanoilla.

Liitetiedostojen esikatselu: Liitetiedostojen esikatselulla tarkoitetaan ohjelman sisäistä toimintoa, jolla liitetiedostoja voidaan tarkastella ilman niiden lataamista käyttäjän koneelle.

Sähköpostilaatikon tilan tarkastus: Sähköpostilaatikon tilalla tarkoitetaan käyttäjälle määritettyä tallennustilaa. Tarkastaminen tarkoittaa siis yksinkertaisesti sitä että käyttäjä näkee itse vapaana olevan tilan ilman että sitä tarvitsee kysyä ylläpitäjältä. Tämä säästää palveluntarjoajalta paljon aikaa, sillä täyteen tulleet laatikot ja kyselyt laatikon tilasta kuormittavat palveluntarjoajan asiakaspalvelua.

Viestien salaaminen: Viestien salaamisella tarkoitetaan viestien salaamista lähetysvaiheessa jollain tunnetuista salaamenetelmistä.

Näkymä sähköpostiketjuille: Näkymä sähköpostiviestiketjuille tarkoittaa näkymää, jossa saman otsikon alle tulleet sähköpostit niputetaan yhdeksi keskusteluksi.

Tuki IMAP -kansioille: Tuki IMAP -kansioille on yksi IMAP -protokollan mahdollistama ominaisuus, jossa käyttäjä voi jakaa sekä hallinnoida useita sähköpostilaatikoita kansiona.

Käyttäjän kaksivaiheinen autentikointi: Käyttäjän kaksivaiheinen autentikointi tarkoittaa nimensä mukaisesti toista varmentavaa autentikointitapaa. Tämä voi esimerkiksi olla koodi, joka lähetetään käyttäjän puhelimeen ja jonka tämä joutuu kirjautuessaan antamaan. Tämä voi aiheuttaa palveluntarjoajalle tarpeen lisälaitteistolle ja siten voi olla palveluntarjoajan näkökulmasta myös huono ominaisuus, mutta jotkin käyttäjät voivat haluta tällaisen ominaisuuden.

Varmuuskopiointi: Varmuuskopiointi tarkoittaa viestien tai käyttäjän määrittämien tiedostojen varmuuskopiointia erilliselle palvelimelle.

Pikanäppäimet: Pikanäppäimet ovat käyttäjän joko itsensä määrittelemiä tai ohjelmassa valmiina olevia pikakomentoja eri toiminnoille.

8.2 Vertailu eri alustojen välillä

Google -hakukoneella löytyi seuraavat sähköpostialustat:

Horde Webmail (URL: <https://www.horde.org/apps/webmail>)

Rainloop (URL: <http://www.rainloop.net/>)

SquirrelMail (URL: <http://squirrelmail.org/index.php>)

Zimbra (URL: <https://www.zimbra.com/>)

Roundcube (URL: <https://roundcube.net/>)

Kopano (URL: <https://kopano.com/>)

Näiden vertailuksi on tehty alla oleva taulukko, johon on valittu alustoissa mainittuja ominaisuuksia.

Taulukko 2. Vertailu sähköpostialustoista. K = Kyllä. E = Ei. T = Tulossa. Ilman merkintää olevista ominaisuuksista ei ole tietoa.

	Horde Webmail	Rainloop	Squirrel-Mail	Zimbra CE	Roundcube	Kopano
Ilmainen	K	K	K	E	K	E
Selainpohjainen	K	K	K	K		K
Mobiilituki				K	T	K
POP3 -tuki	K			K		
IMAP -tuki	K	K	K	K	K	
Sähköpostien suodatus	K	K				K
Sähköposteista hakeminen	K			K	K	K
Liitetiedostojen esikatselu	K	K			T	K
Sähköpostilaatikon tilan tarkastus	K					
Viestien salaaminen	K				K	
Näkymä sähköpostiketjuille	K			K	K	
Tuki IMAP -kansioille	K	K	K	K	K	K
Käyttäjän kaksivaiheinen autentikointi				K		
Varmuskopiointi				K		
Pikanäppäimet	K	K			T	

Taulukosta on nähtävissä että selkeästi monipuolisin alustoista on ensiksi mainittu Horden Webmail. Tuloksessa on kuitenkin huomioitava mahdolliset muut kriteerit, joita erilaisilla palveluntarjoajilla voi olla. Tärkeää on myös huomata ettei tutkimus ota kantaa tietoturvaan ja sen osalta alustoissa voi olla eroja jotka palveluntarjoajan on syytä ottaa huomioon. Palveluntarjoajan tarpeet voivat myös vaihdella jolloin jokin toinen taulukossa esitellyistä alustoista voi olla paremmin käyttöön soveltuva. Joitain ominaisuuksia olisi parempi arvostella syvällisemmin, kuin pelkästään Kyllä/Ei -asteikolla, mutta tämän tutkimuksen laajuudessa se ei ole mahdollista, sillä se vaatisi eri alustojen asentamisen sekä testaamisen.

9. Yhteenveto

Aikaisemmat tutkimukset keskittyvät enemmän käyttäjien kuin organisaation näkökulmaan. Tärkeä huomio POP3- ja IMAP-protokollia verrattaessa olikin palveluntarjoajan näkökulma, sillä kuluttajan kannalta IMAP-protokolla on selvästi parempi, mutta sen sijaan palveluntarjoajalle se on selvästi huonompi vaihtoehto. Nykypäivänä palveluntarjoajan tulisi kuitenkin tukea molempia protokollia ja siten tutkimuksen pääpaino siirtyi protokollien tutkimiseen sekä sopivan sähköpostialustan löytämiseen. Sähköpostialustoista selkeästi monipuolisin ja siten varmin valinta on Horden Webmail, joka sattuu myös olemaan ilmainen. Vertailussa ei otettu kantaa eri alustojen ulkonäköön, mikä voi olla joillekin palveluntarjoajille tärkeä kriteeri esimerkiksi brändin takia.

Tästä tutkimuksesta rajoitettiin tietoturva pois sen ollessa liian suuri kokonaisuus tämän kaltaiselle tutkimukselle. Jatkotutkimuksia olisikin hyvä tehdä tietoturva huomioiden sekä mahdollisesti tehokkuutta mitaten eri alustojen välillä. Myös eri alustojen vaatimukset palvelimen tehojen suhteen olisi hyvä tietää palveluntarjoajan näkökulmasta ja lopullista vertailua tehdessä.

Lähdeluettelo

- Dean, T. (2009). *Network+ Guide to Networks*. Boston, MA: Course Technology Press
- Karvinen, M. 1997: SMTP, Internet WWW-sivu, Viitattu 8/2017 Saatavilla <http://www.tml.hut.fi/Studies/Tik-110.300/1998/Essays/smtp.html>
- Kissell, J. (2006). The IMAP advantage. *Macworld*, 23, 75-77.
- Raghubandan, HP. (2009). Autonomic E-mail services for better storage management.
- RFC1939. Post Office Protocol – Version 3. Viitattu 1/2017 Saatavilla <https://tools.ietf.org/html/rfc1939>
- RFC2045. Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies Viitattu 2/2018 Saatavilla <https://tools.ietf.org/html/rfc2045>
- RFC2046. Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types. Viitattu 2/2018 Saatavilla <https://tools.ietf.org/html/rfc2046>
- RFC2449. POP3 Extension Mechanism. Viitattu 8/2017 Saatavilla <https://tools.ietf.org/html/rfc2449>
- RFC2505. Anti-Spam Recommendations for SMTP MTAs. Viitattu 12/2017 Saatavilla <https://www.ietf.org/rfc/rfc2505.txt>
- RFC2683. IMAP4 Implementation Recommendations Viitattu 6/2017 Saatavilla <https://tools.ietf.org/html/rfc2683>
- RFC3501. Internet Message Access Protocol – Version 4rev1. Viitattu 1/2017 Saatavilla <https://tools.ietf.org/html/rfc3501>
- RFC5068. Email Submission Operations: Access and Accountability Requirements Viitattu 2/2018 Saatavilla <https://tools.ietf.org/html/rfc5068>
- RFC5321. Simple Mail Transfer Protocol Viitattu 2/2018 Saatavilla <https://tools.ietf.org/html/rfc5321>
- RFC5322. Internet Message Format. Viitattu 2/2018 Saatavilla <https://tools.ietf.org/html/rfc5322>
- RFC5598. Internet Mail Architecture Viitattu 12/2017 Saatavilla <https://tools.ietf.org/html/rfc5598>
- Srinivasan, J. (2011). EMFS: Email-based Personal Cloud Storage.
- Tzerefos P., Smythe C., Stergiou I. and Cvetkovic S., "A comparative study of Simple Mail Transfer Protocol (SMTP), Post Office Protocol (POP) and X.400 Electronic Mail Protocols," *Proceedings of 22nd Annual Conference on Local Computer Networks*, Minneapolis, MN, 1997, pp. 545-554.