

2. Sovelluksia ja sovellusprotokollia

- 2.1. WWW (World Wide Web)
- 2.2. Sähköposti
- 2.3. DNS (Domain Name System)

1/23/2004

1

Prosessien etäkommunikointi

- Eri koneissa olevien prosessien kommunikointia lähettämällä sanomia toisilleen
 - KJ huolehtii saman koneen prosessien kommunikoinnista
- **Pistoke (socket)**
 - prosessi kirjoittaa verkkoon ja lukee verkosta samalla tavoin kuin kirjoittaa tiedostoon ja lukee tiedostosta
 - API (application programmers' interface)
 - verkkosovelluksen ohjelmointirajapinta

1/23/2004

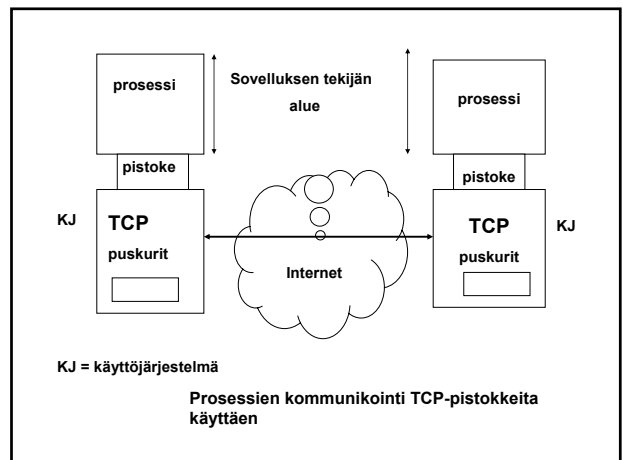
4

Verkkosovellus \Leftrightarrow sovellusprotokolla

- Sovellusprotokolla on vain osa hajautettua sovellusta
- Esim. **WWW**
 - selain, **www-palvelin**, dokumentin rakenne (**HTML**) ja sovellusprotokolla (**HTTP**)
 - **HTTP**
 - selaimen ja verkkopalvelimen kommunikointiin

1/23/2004

2



Sovellusprotokolla määrittelee

- käytetyt **sanomatyytit**
 - pyyntö (request)
 - vastaus (response)
- **sanomien rakenteen** (syntaksi)
 - mitä kenttiä, minkä kokoisia ja missä kohtaa
- kenttien merkityksen (semantiikka)
- 'säännöt', milloin mikin sanoma lähetetään

1/23/2004

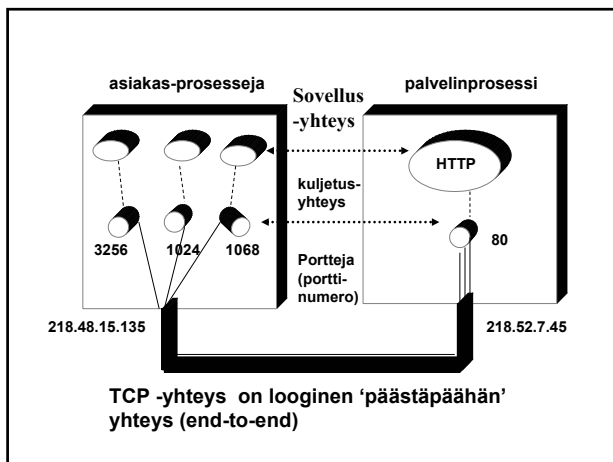
3

Osoittaminen

- **IP-osoite => oikea kone**
 - koneen yksilöivä tunniste
 - koneen verkkoliitännän yksilöivä tunniste
 - verkko-osa osoitteesta yksilöi verkon
 - koneosa yksilöi koneen verkossa
- **Porttinumero => oikea prosessi**
 - yleisillä palveluilla standardoidut porttinumerot
 - **www-palvelin** kuuntelee porttia 80
 - **postipalvelin** kuuntelee porttia 25

1/23/2004

6



Internetin kuljetusprotokollat

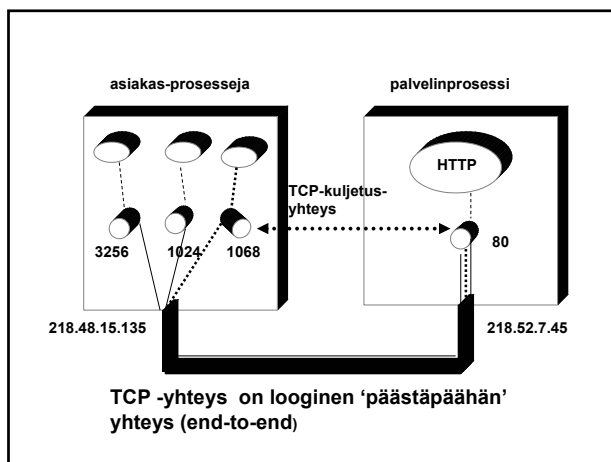
- **TCP (Transmission Control Protocol)**
 - yhteydellinen palvelu
 - yhteyden muodostus ennen datan siirtoa
 - kättely (handshaking)
 - kaksisuuntainen TCP-yhteys
 - yhteyden purku
 - luotettava kuljetuspalvelu
 - ruuhkanhallinta
 - ei takuita siirtonopeudelle eikä viiveelle

1/23/2004 10

Käyttäjänedustaja (User Agent)

- **Käyttäjän ja verkkosovelluksen rajapinta**
 - verkossa selain
 - hakee sivun ja näyttää sen
 - käynnistää Java-sovelmat
 - asiakkaan puolen HTTP-protokolla lähettää ja vastaanottaa sanomia piskokkeen kautta
 - Netscape Navigator, Internet Explorer
 - sähköpostissa postiohjelma ('mail reader')
 - graafinen käyttöliittymä sanomien laatimiseen ja lukemiseen
 - asiakkaan puolen SMTP sanomien lähettämiseen ja esim. POP tai IMAP sanomien hakemiseen
 - Eudora, Netscape Communicato, Microsoft Outlook, pine

1/23/2004 8



Sovellus ja kuljetuspalvelun laatu

- **virheettömyys**
 - pankkisovellus
 - videoneuvottelu
 - puhelu
- **kapasiteetti eli minimisiirtonopeus**
 - Internet-puhelin 32 Kbps, video 10Kbps-5 Mbps
 - sähköposti
- **tosiakaisuus**
 - usean pelaajan taistelupeli: viive korkeintaan 100 ms
 - dokumentin siirto

1/23/2004 9

- **UDP (User Datagram Protocol)**
 - minimaalinen kuljetuspalvelu
 - ei yhteyden muodostusta eikä purkua
 - sanoma vain lähetetään verkkoon
 - ei mitään takuita sanoman perillemenosta
 - saapuneet sanomat voivat olla epäjärjestyksessä
 - ei ruuhkanvalvontaa
 - lähettäjä voi lähettää UDP-pistokkeeseen niin paljon kuin haluaa

1/23/2004 12

2.1. WWW

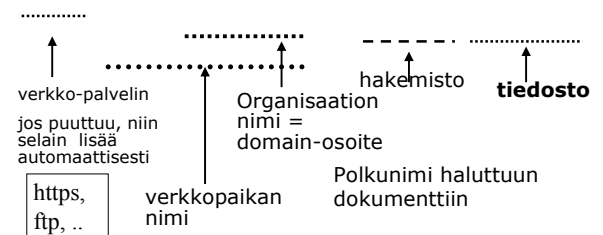
- WWW on joukko yhteentoimivia palvelimia ja asiakkaita, jotka puhuvat samaa HTTP-kieltä (-protokollaa)
- graafinen asiakasohjelma selain
 - Netscape, Explorer, Mozilla, ...
- kykenee 'avaamaan' URL-linkin = muodostamaan TCP-yhteyden ko. verkkopalvelimeen ja hakemaan sieltä ko. tiedoston

1/23/2004

13

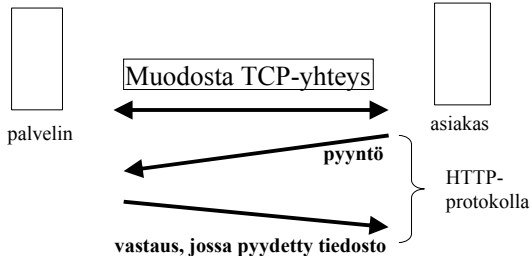
URL (Uniform Resource Locator)

<http://www.nmib.com/glossary.index.html>



<http://www.cs.princeton.edu/index.html>

1. Muodostetaan TCP- yhteys koneeseen: www.cs.princeton.edu
2. Pyydetään sieltä tiedosto käyttäen HTTP-protokollaa: [index.html](http://www.cs.princeton.edu/index.html)



Muita URLeja:

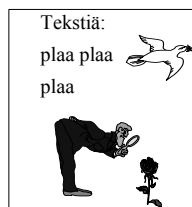
- <ftp://usc.edu/pubs/myfile.doc>
 - **ftp-palvelulla haettava tiedosto**
- <news:uk.finance>
 - **uutisryhmä**
- <file:///C:/webs/html/mottle.gif>
 - tiedosto haetaan käyttöjärjestelmän avulla (ei käyttäen http:tä)

1/23/2004

17

HTTP (Hypertext Transfer Protocol)

- Protokolla asiakasohjelman ja palvelinohjelman väliseen kommunikointiin
(HTTP/1.1: <http://www.w3.org/Protocols/rfc2616/rfc2616.html>)
 - **www-sivujen hakemiseen**
 - HTML-tiedosto, jossa
 - HTML-tiedostoja
 - JPEG-kuva, GIF-kuva
 - Java-sovelma
 - ..
 - **sivuilla URL-osoite**
 - tilaton protokolla



1/23/2004

5

HTTP-asiakas

- selain (= asiakas) pyytää verkkosivua
 - **muodostaa TCP-yhteyden** palvelinkoneeseen
 - DNS auttaa domain-nimen muuttamisessa IP-osoitteeksi
 - **WWW-palvelimen kuuntelema portti on 80**
 - **lähettää HTTP-pyyntösanoman** TCP-yhteyteen liitettyyn pistokkeeseen
 - **vastaaottaa palvelimen lähettämän vastausanoman**, jossa on sisällä pyydetty sivu
 - tämän jälkeen TCP-yhteys suljetaan
 - selain tutkii sivua ja näkee uudet viitteet kuviin, jotka se sitten hakee samalla tavalla
 - selain näyttää pyydetyn sivun käyttäjälle

1/23/2004

18

Palvelinprosessi

- Kuuntelee TCP-porttia 80 yhteyspyyntöjä varten
 - pyyntö yleensä selaimelta
 - kun käyttäjä on ilmoittanut haluavansa ko. dokumentin
 - esim. klikkaamalla tekstissä olevaa URL-linkkiä
 - tai antamalla URL-linkin yhteyspyynnössä
 - tai selain itse huomaa linkin, josta pitää hakea

1/23/2004

19

Suorituskyky?

- Jos haetaan 10 objektaa
 - 10 TCP-yhteyden muodostusta ja purkua
 - kukin 2 sanomaa = RTT (round-trip time) => 20 RTT
 - hidas aloitus (slow start) hidastaa lähety nopeutta
 - voidaan avata useita rinnakkaisia yhteyksiä
 - puskuritilat yhteyksille
- käytetään pysyvää TCP-yhteyttä
 - palvelin jättää yhteyden sulkematta
 - muut pyynnöt ja vastaukset käyttävät samaa yhteyttä

1/23/2004

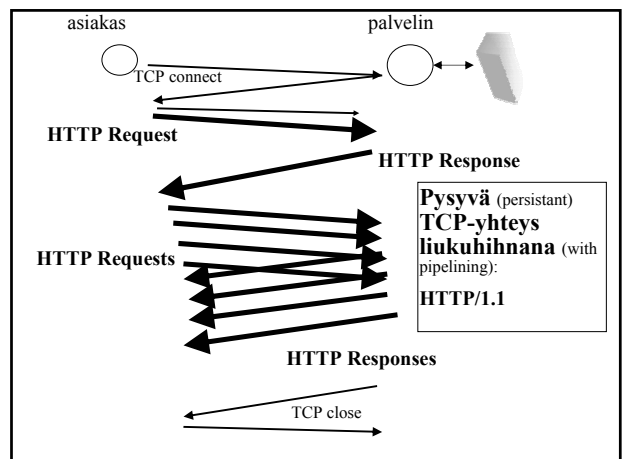
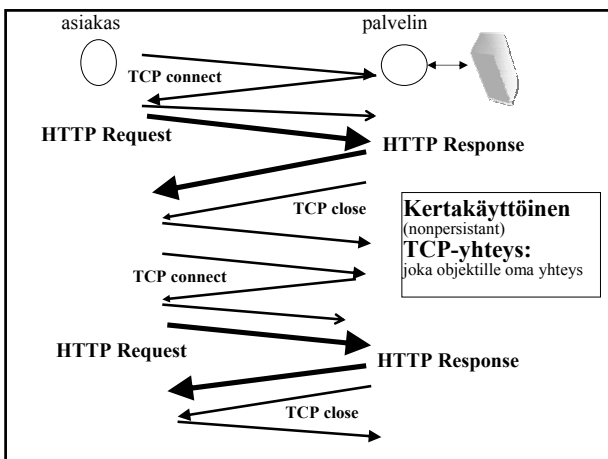
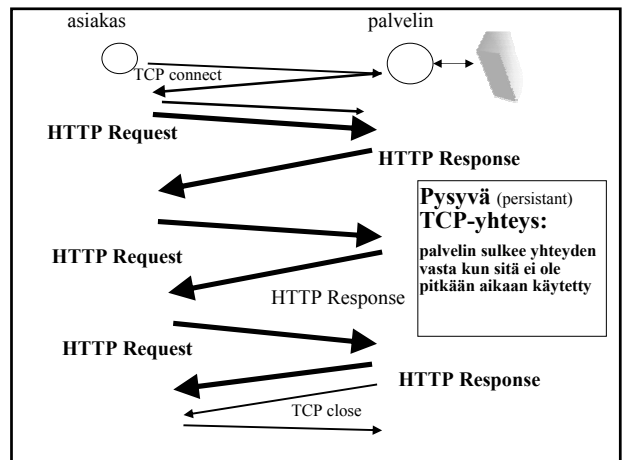
22

HTTP-pyyntön tullessa

- verkkopalvelin
 - vastaanottaa pyyntösanomaa,
 - hakee pyydetyn sivun tai objektin omasta muististaan,
 - liittää sen HTTP-vastaussanomaa ja
 - lähettää sitä pyytäneelle selaimelle
 - pyytää TCP-yhteyden sulkemista

1/23/2004

20



HTTP-sanomat (esimerkki)

■ HTTP Request Message:

GET /jokuhakemisto/sivu.html HTTP/1.1

Host: WWW.jokupaikka.fi

Connection: close

User-agent: Mozilla/4.0

Accept-language: fi

(ylimääräinen CR ('control') ja LF ('line feed'))

1/23/2004

25

HTTP-sanomat (esimerkki)

■ HTTP Response Message:

HTTP/1.1 200 OK

Connection: close

Date: Thu, 19 Oct 2000 12:00:15 GMT

Server: Apache/1.3.0 (Unix)

Last-Modified: Mon, 22 Jun 2000 09:23:24 GMT

Content-Length: 6821

Content-Type: text/html

(data data data data data ...)

1/23/2004

28

Pyyntösanomien yleinen rakenne

GET /jokuhakemisto/sivu.html HTTP/1.1

metodi	SP	URL	SP	versio	CR	LF
--------	----	-----	----	--------	----	----

Otsakekenttä : kentän arvo CR LF

SP='space'
eli välilyönti

... Lisää otsakerivejä

Otsakekenttä : kentän arvo CR LF

CR + LF =
rivin
päättäminen

CR LF

Runko-osa
käytössä esim. POST-metodissa

Vastausanomien yleinen rakenne

versio	SP	statuskoodi	SP	fraasi	CR	LF
--------	----	-------------	----	--------	----	----

Otsakekenttä : kentän arvo CR LF

... Lisää otsakerivejä

Otsakekenttä : kentän arvo CR LF

CR LF

Runko-osa

Otsakekenttä : kentän arvo	CR	LF
----------------------------	----	----

Host: WWW.jokupaikka.fi kone, jossa dokumentti on

Connection: close sulje yhteys lähetyksen jälkeen

User-agent: Mozilla/4.0 selainen tyyppi

Accept-language: fi dokumentin kieli

Pyyntömetodeja mm.

- GET sivun lukeminen
 - aina tai vain tietyn päiväyksen jälkeen muutettu sivu
- HEAD pelkän otsikon lukeminen
- PUT sivun tallettaminen
- POST lisäys verkkosivulle
- DELETE sivun poistaminen

1/23/2004

30

Statuskoodeja ja fraaseja

- 200 OK
 - pyyntö onnistui , pyydetty sivu vastauksessa
- 301 Moved Permanently
 - uusi URL on otsakekentässä Location
- 400 Bad Request
- 404 Not Found
- 505 HTTP Version Not Supported

1/23/2004

31

Evästeitä:

(Windows: C:\Documents and Settings\Windows Explorer: C:\Program Files\ICookies, Linux: ~/.netscape/cookies.)

www.download.com lähettäjän osoite
FALSE / FALSE turvayhteyksien käyttö
946598400 vanhenemispäivä
csr 'piparin' nimi
/=/PC/FrontDoor/0^1^0- 0^01.html data


www.harmony-central.com FALSE / FALSE
1034625738 s 13013000876945738607

www.illuminatus.com TRUE / FALSE
945734399 Count 3

1/23/2004

34

Käyttäjän tunnistaminen

- Autentikointi
 - omat koodit ja otsakkeet tätä varten
 - 401 Authorization Required -statuskoodi
 - WWW-Authenticate -otsake kertoo kuinka
 - asiakas lähettää autentikointitiedot joka pyynnössä (Authorization- otsakerivi)
- Eväste (pipari, cookie) 
 - asiakkaalle talteen käyttäjäkohtainen tunnuskoodi, joka esitetään joka pyynnön yhteydessä

1/23/2004

32

Verkkovälimuisti

- Säilyttää kopioita haetuista sivuista
 - viimeksi haetut sivut
- pyyntö ohjautuu ensin välimuistiin
 - haetaan verkon yli vasta, jos ei löydy välimuistista
- etuja:
 - lyhentää kyselyaikaa
 - vähentää Internetin liikennettä
 - tuo nopean jakelukanavan hitaille palvelimille

1/23/2004

35

asiakas ○ ○ palvelin



Eväste-
tiedosto

Käyttö:

- autentikointi ('yksilöi käyttäjän')
- kohdistettu mainonta ('tietoja käyttäytymisestä => 'oikea mainos')
- virtuaalinen ostoskärri

palvelin ○



Eväste-
profiilit

Set-cookie: 167845<CR><LF>

Cookie: 167845<CR><LF>

Cookie: 167845<CR><LF>

palvelin

64 kbps

Internet

1.5 Mbps

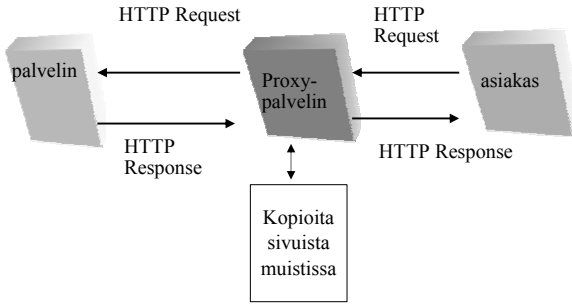
Proxy-
palvelin

Yrityksen
LAN-
verkko

Välimuistin eli ns.
Proxy-palvelimen
käyttö

10-100 Mbps

Haku palvelimelta vain, jos ei ole jo muistissa



2.2. Sähköposti

- Käyttäjiliitäntä (user agent)
 - sanomien kirjoittaminen, lukeminen ja lähettäminen
- Postipalvelin (mail server)
 - postilaatikot
- Postiprotokolla SMTP
 - sanomien toimittaminen lähettäjän postijärjestelmästä vastaanottajan postijärjestelmään
 - sanoman muoto
 - tekstisanoma, MIME-sanoma

1/23/2004

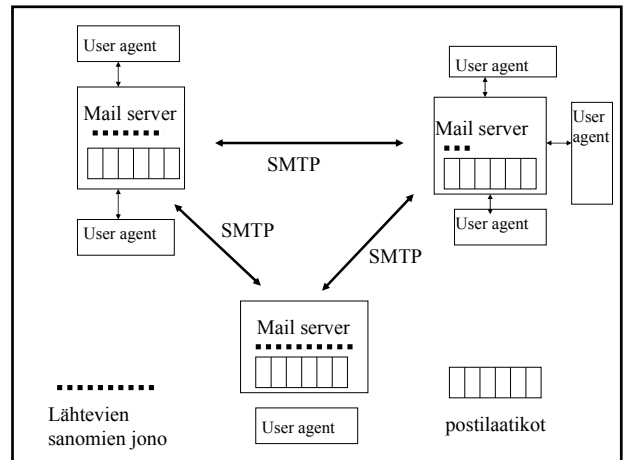
40

Conditional GET

- Välimuistiin talletettu objekti haetaan verkosta, jos objekti on muutettu
 - tallessa vanha versio
- Get
 - If-modified-since: Mon, 22 Jun 2000 09:23:24**
- HTTP/1.0 304 Not Modified

1/23/2004

38



HTML (HyperText Markup Language)

- SGML (Standard Generalized Markup Language)
 - merkkaukieli
 - kertoo, kuinka dokumentit muotoillaan
 - TeX, troff,
 - ladontamerkinnot
 - WYSIWYG
- XML (Extensible Markup Language)
- Näistä enemmän:

582302 Digitaalisen median tekniikat (2ov)
582304 XML-metakieli (2ov)

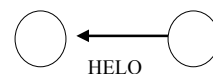
1/23/2004

39

SMTP (Simple Mail Transfer Protocol)

- Postipalvelimet käyttävät SMTP-protokollaa sanomien siirtoon
 - sanomat lähetävä postipalvelu muodostaa TCP-yhteyden vastaanottajan postipalvelimeen
 - 'kättelyjen' jälkeen siirretään sanomat

Vastaanottava on palvelin



Lähetävä on asiakas

1/23/2004

42

Lähetävä postipalvelin

- muodostaa TCP-yhteyden
- esittelee itsensä
- kertoo lähettäjän ja vastaanottajan
- lähettää sanoman, jos vastaanottava palvelin antaa luvan

Vastaanottava postipalvelin

- kuuntelee TCP-porttia **25**
- hyväksyy yhteydenmuodostuspyynnöt
- vastaanottaa ja siirtää sanomat postilaatikoihin
- palauttaa virheilmoitukset

```
S: 220 helsinki.fi
C: HELO princeton.edu
S: 250 Hello princeton.edu
C: MAIL FROM: <Bob@princeton.edu>
.....
S: 250 Message accepted for delivery
C: HELO princeton.edu
S: 250 Hello princeton.edu
C: MAIL FROM: <Alice@princeton.edu>
.....
S: 250 Message accepted for delivery
.....
C: QUIT
S: 221 princeton.edu closing connection
```

1/23/2004

46

```
S: 220 helsinki.fi
C: HELO princeton.edu
S: 250 Hello princeton.edu
C: MAIL FROM: <Bob@princeton.edu>
S: 250 <Bob@princeton.edu> OK
C: RCPT TO: <pekka.puupaa@cs.helsinki.fi>
S: 250 <pekka.puupaa@cs.helsinki.fi> OK
C: DATA
S: 354 Enter mail, end with "." on a line by itself
C: dataa ... dataa
C: dataa ... dataa
C: .
S: 250 Message accepted for delivery
C: QUIT
S: 221 princeton.edu closing connection
```

1/23/2004

44

Push- ja pull-protokollat

- työntöprotokolla (PUSH)
 - asiakas lähettää tiedoston
 - SMTP
- noutoprotokolla (PULL)
 - asiakas pyytää tiedostoa
 - HTTP
 - poikkeuksena PUT- ja POST-metodit

1/23/2004

47

- TCP-yhteys on pysyvä
 - yhteydellä voidaan lähettää useita sanomia
- Yhteydellä voidaan lähettää vain ASCII-sanomia

- sanomatekstin lopettaa omalla rivillään irrallisena oleva piste (".") (*Tämä on vanha protokolla!*)

eli ASCII-merkit **CRLF.CRLF**

- jos sanomassa on binääridataa, esim. kuvia, niin siellä voisi esiintyä irrallinen piste eli CRLF.CRLF => sanoma katkeaa. => binääridata koodataan ASCIIksi siten, ettei siinä esiinny CRLF.CRLF : iä

1/23/2004

45

Sähköpostisanoman muoto

To:
CC:
From:
Message-Id:
Received:
Date:
Reply-To:
Subject:

Yleisiä sanoman otsakekenttiä
kentät erotettu rivinvaihdolla

RFC 822 sallii myös omien kenttien käytön

1/23/2004

48

MIME (Multipurpose Internet Mail Extension)

- Alunperin vain tekstisanomia (ASCII)
- nykyisin myös
 - erityyppisiä tekstisanomia
 - Word-dokumentteja
 - kiinankielistä tekstiä
 - ääntä,
 - videota
 - kuvia

1/23/2004

49

MIME-sisältötyyppejä

- text/plain; charset=us-ascii
- text/html
- image/gif, image/jpeg, video/mpeg
- application/postscript, application/msword, application/octetstream
- **multipart/mixed**

1/23/2004

52

MIME-laajennus

- Lisää kenttiä otsikkoon

MIME-versio:

Content-Transfer-Encoding:

Content-Type:

1/23/2004

50

Content-Type:
multipart/mixed;Boundary=StartOfNextPart

-- StartOfNextPart

Hei Allu,

sinulle kaunis kuva kissastani Villestä.

-- StartOfNextPart

Content-Transfer-Encoding: base64

Content-Type: image/jpeg

base64 encoded data

.....

.....

....

.....base64 encoded data

-- StartOfNextPart

Haluatko muita kuvia!

1/23/2004

53

From: alice@crepes.fr
To: bob@hamburger.edu
Subject: Picture of yummy crepe
MIME-Version: 1.0
Content-Transfer-Encoding: base64
Content-Type: image/jpeg
(base64 encoded data
.....
..... base64 encoded data)
.

- tyyppien koodaukset siirtoa varten
 - sähköpostisanomassa saa olla vain ASCII-merkkejä
 - ei esimerkiksi ääkkösiä

• **base64-koodaus**

1/23/2004

54

Base64-koodaus

- Sanoman 24 bitin ryhmät jaetaan 6 bitin osiksi, jotka kukin koodataan ASCII-merkkiksi => 64 eri vaihtoehtoa

- 0-25 = A-Z,
- 26-51 = a-z,
- 52-61 = 0-9,
- 62= +, 63 =/

- **Tekstisanoma voidaan koodata myös 7-bitin ASCIIlla.**

1/23/2004

55

Postiyhdyskäytävä (Mail Gateway)

- Yleensä postipalvelin ottaa suoraan yhteyttä vastaanottajan postipalvelimeen
- välissä voi olla postiyhdyskäytäviä
 - tallettavat ja toimittavat eteenpäin sanomia
 - yrittävät lähettämistä useita päiviä
- **Miksi tarvitaan?**
 - ohjaa oikeaan koneeseen: liisa.martinen@helsinki.fi
 - vastaanottajan kone usein pois päältä

1/23/2004

58

Postinhakuprotokollat (mail access protocols)

- Käyttäjän postiohjelma siirtää postinhakuprotokollan avulla sanomat postipalvelimen postilaatikoista käyttäjän PC:lle
 - POP3: yksinkertainen ohjelma, joka siirtää sanomat vastaanottajan PC:lle
 - IMAP: monipuolinen, käyttäjä voi käsitellä postipalvelimen postilaatikkoo kuin se olisi omassa koneessa
- Entä lähetettäessä?

1/23/2004

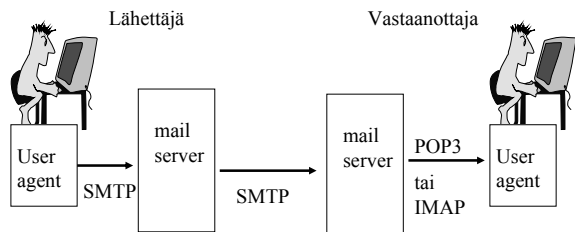
56

2.3. DNS (Domain Name System)

- Internetin hakemistopalvelu
- **Ihmisten ja ohjelmien käyttämien merkkijonomuotoisten nimien muuttaminen binäärisiksi IP-verkko-osoiteiksi**
 - hydra.cs.helsinki.fi => 218.214.4.29
- **hierarkkinen nimentätapa + hajautettu hakemisto IP-osoitteiden selvittämiseksi**
 - esim. sähköpostiosoitteita, koneiden IP-osoitteita

1/23/2004

59



Koska SMTP on 'PUSH'-protokolla, sitä ei voi käyttää sanomia haettaessa ('PULL').

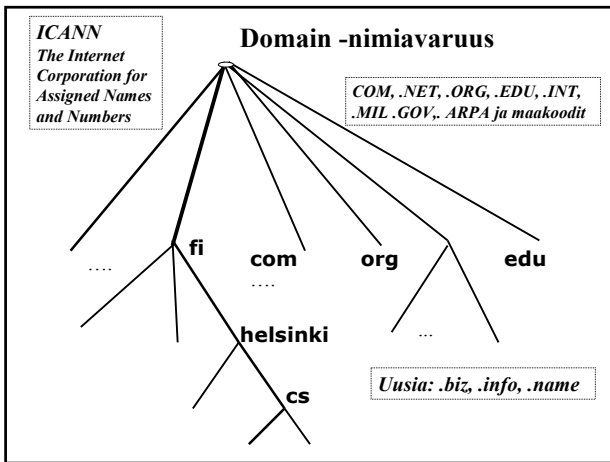
POP3, IMAP tai yhä yleisemmin käytetään selainta ja HTTP-protokollaa

DNS- nimiavaruus

- **Domain (nimihallinta-alue)**
 - maiden tunnukset
 - fi, fr, de, uk, ...
 - yleiset (yleensä USA)
 - com, edu, gov, int, mil, net, org
 - esimerkkejä:
 - cs.helsinki.fi
 - www.cnn.com
 - www-dept.cs.ucl.ac.uk

1/23/2004

60



■ Kysely ensin paikalliselle nimipalvelimelle

- jos nimi on sen vastuulla => authoritative record
- jos sillä ei ole mitään tietoa nimestä, se lähettää kyselyn ylimmän tason ns. juurinimipalvelijalle
 - (DNS root server, <http://www.wia.org/pub/rootserv.html>, <ftp://ftp.rs.internic.net/domain/named.root>)
 - esim. Kysely "linda.cs.yale.edu" lähetetään nimipalvelimelle "edu-server.net", joka vastaa vyöhykkeestä "edu".

1/23/2004 64

IP-nimen selvittäminen

- sovellusohjelma kutsuu kirjastorutiinia parametrina nimi merkkijonona
 - esim Unix:ssa `gethostbyname()`
- kirjastorutiini lähettää UDP-datasähkeen paikalliselle DNS-palvelimelle, joka etsii nimeä vastaavan IP-osoitteen ja palauttaa sen kirjastorutiinille
 - etsinnässä tarvitaan usein monien palvelimien yhteistyötä

1/23/2004 62

■ Juurinimipalvelin tietää, mikä nimipalvelin on vastuussa kysytystä nimestä

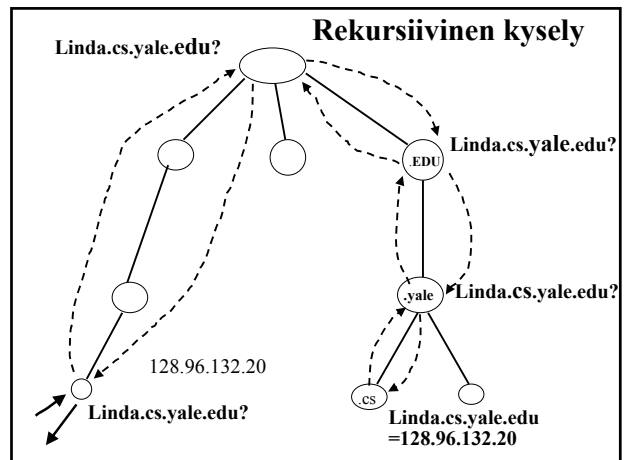
- tämä on tosin voinut jakaa vastuun joillekin toisille nimipalvelimelle
- mutta tietää, mikä niistä on vastuussa juuri kysytystä nimestä

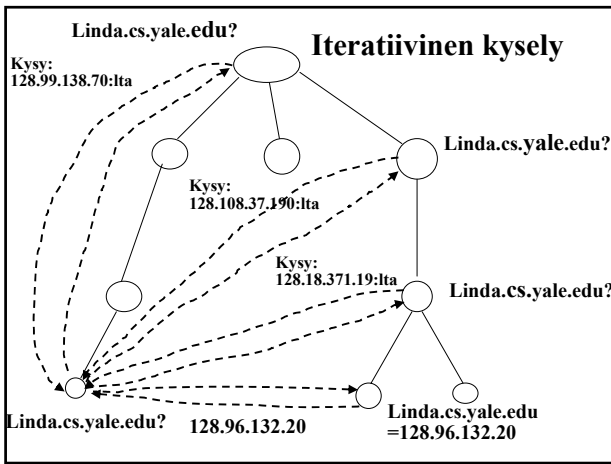
1/23/2004 65

Nimipalvelimien hierarkia

- DNS-nimiavaruus jaettu vyöhykkeisiin (zone)
 - kukin vyöhyke kattaa osan nimipuusta
 - vyöhykkeellä on yksi siitä vastaava nimipalvelija (primary) ja yksi tai useita apunimipalvelijoita (secondary)
- vyökykejako on hallinnollinen
 - tarpeen mukaan nimipalvelijoita vastaamaan omasta alueestaan

1/23/2004 63





DNS-välimuisti

- Suorituskyvyn parantamiseksi nimipalvelijat varastoivat välimuistiinsa DNS-tietueita
- ei tarvitse hakea uudelleen
- elinaika kertoo voimassaoloajan

1/23/2004 70

Hakemistotiedot

- DNS-hakemistotietue (resource record)
 - (nimi, arvo, tyyppi, tietueen elinaika)
 - Tyyppi määrää nimen ja arvon merkityksen
 - A-tyyppi
 - koneen domain nimi , arvo on IP-osoite
 - **NS-tyyppi**
 - domain-nimi, arvo on alueesta vastaavan nimipalvelijan koneen domain-nimi
 - CNAME-tyyppi
 - alias-nimeä vastaava 'kanoninen' nimi
 - MX-tyyppi
 - aliasnimeä vastaava postipalvelin

1/23/2004 68

DNS-sanomat

Identification	Flags
Number of questions	Number of answers
Number of authority RRs	Number of additional RRs
Questions (name and types for queries)	
Answers (RRs in response to queries)	
Authority (RRs for authoritative servers)	
Additional information	

1/23/2004 69