

•
•
•

2. Sovelluksia ja sovellusprotokollia

2.1. WWW (World Wide Web)

2.2. Sähköposti

2.3. DNS (Domain NameSystem)

-
-
-

Verkkosovellus \Leftrightarrow sovellusprotokolla

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

•
•
•

Sovellusprotokolla määrittelee

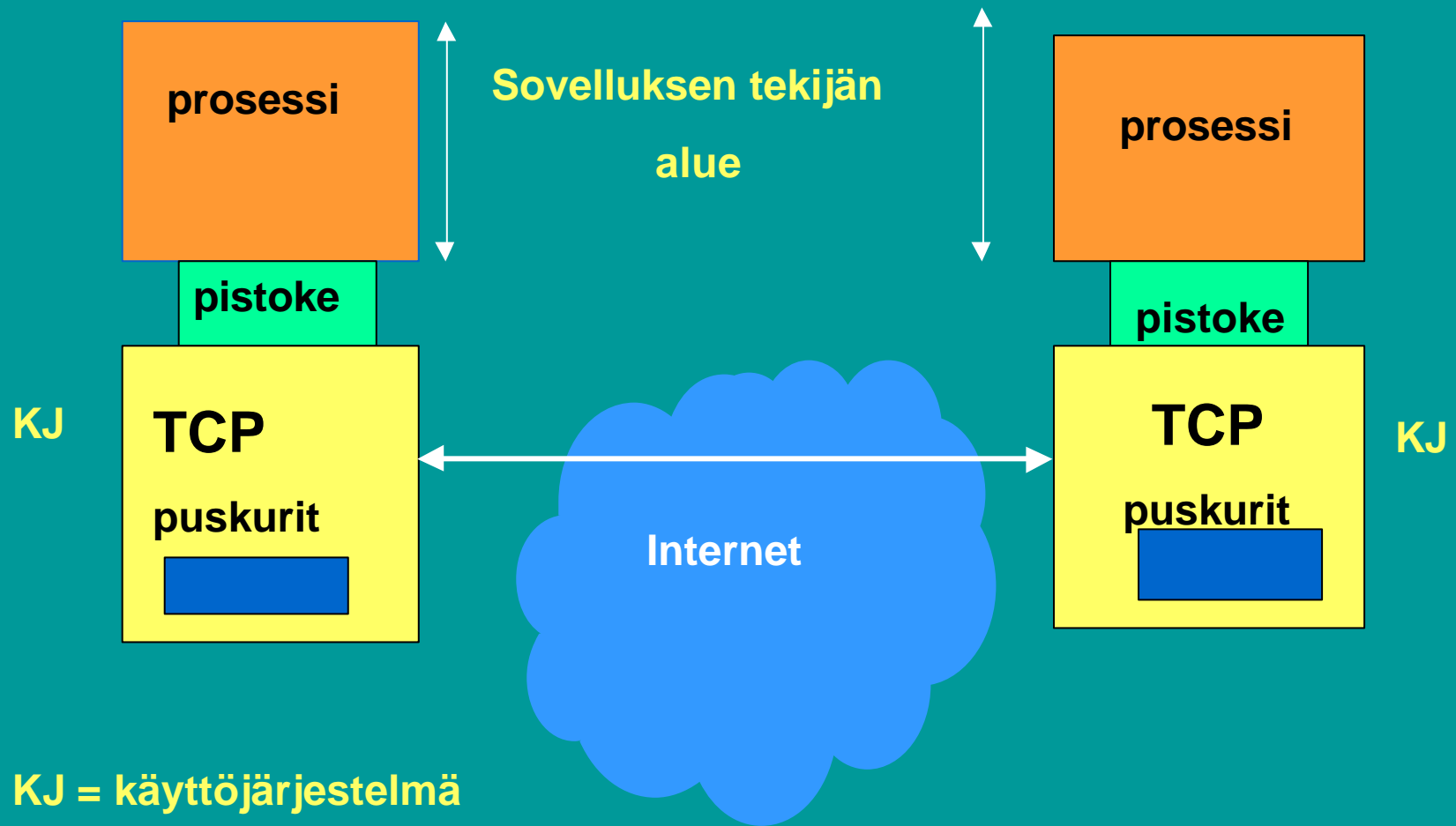
- **käytetyt sanomatyypit**
 - pyyntö (request)
 - vastaus (response)
- **sanomien rakenteen (syntaksi)**
 - mitä kenttiä, minkä kokoisia ja missä kohtaa
- **kenttien merkityksen (semantiikka)**
- **‘säännöt’, milloin mikin sanoma lähetetään**

-
-
-

Prosessien etäkommunikointi

- **Pistoke (socket)**

- prosessi kirjoittaa verkkoon ja lukee verkosta samalla tavoin kuin kirjoittaa tiedostoon ja lukee tiedostosta
- API (application programmers' interface)
 - verkkosovelluksen ohjelmointirajapinta

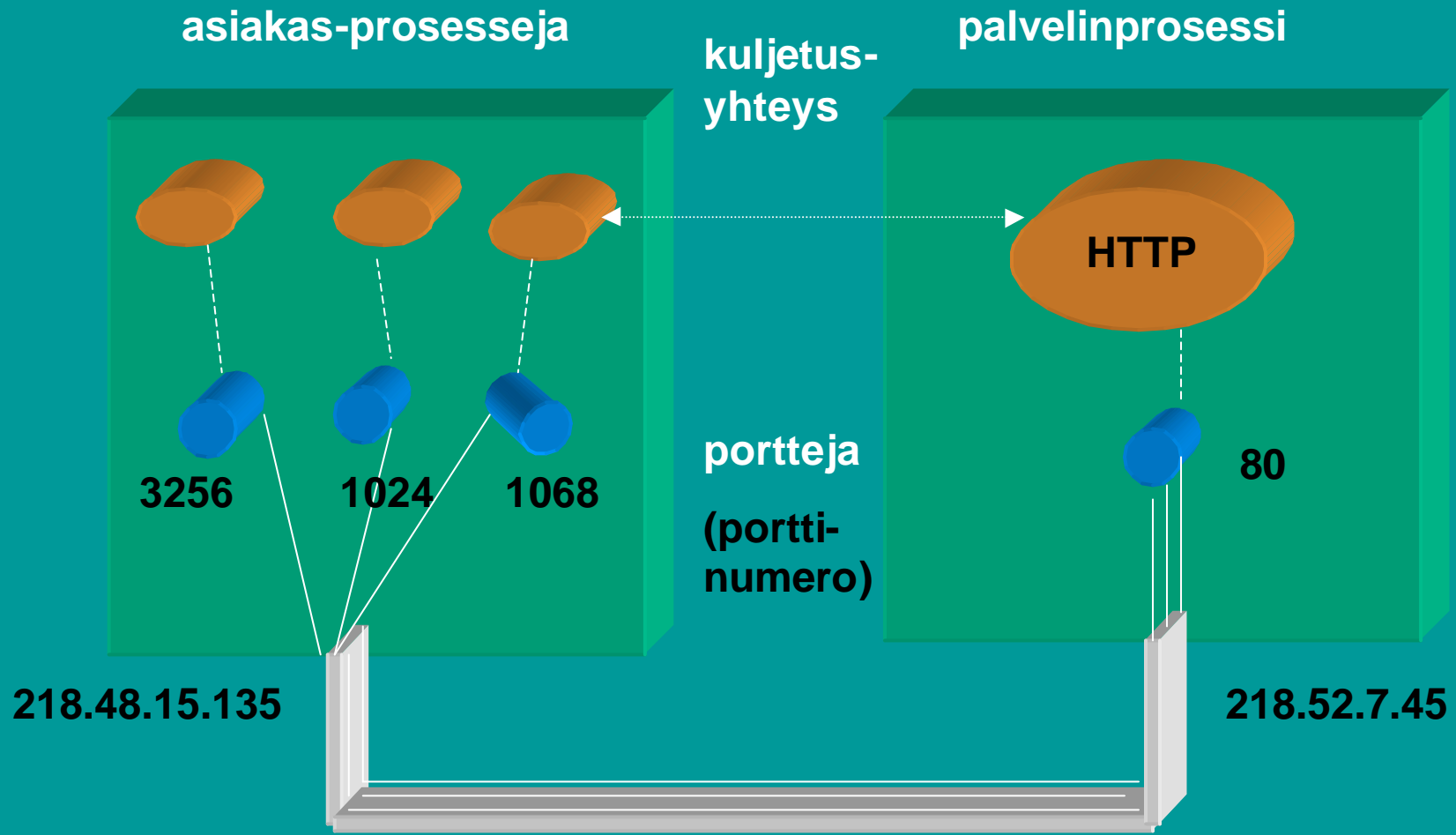


Prosessien kommunikointi TCP-pistokkeita käyttäen

-
-
-

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



TCP -yhteys on looginen 'päästäpäähän' yhteys (end-to-end)

-
-
-

Internetin kuljetusprotokollat

- **TCP**

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

-
-
-

• UDP

- minimaalinen palvelu
- 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

•
•
•

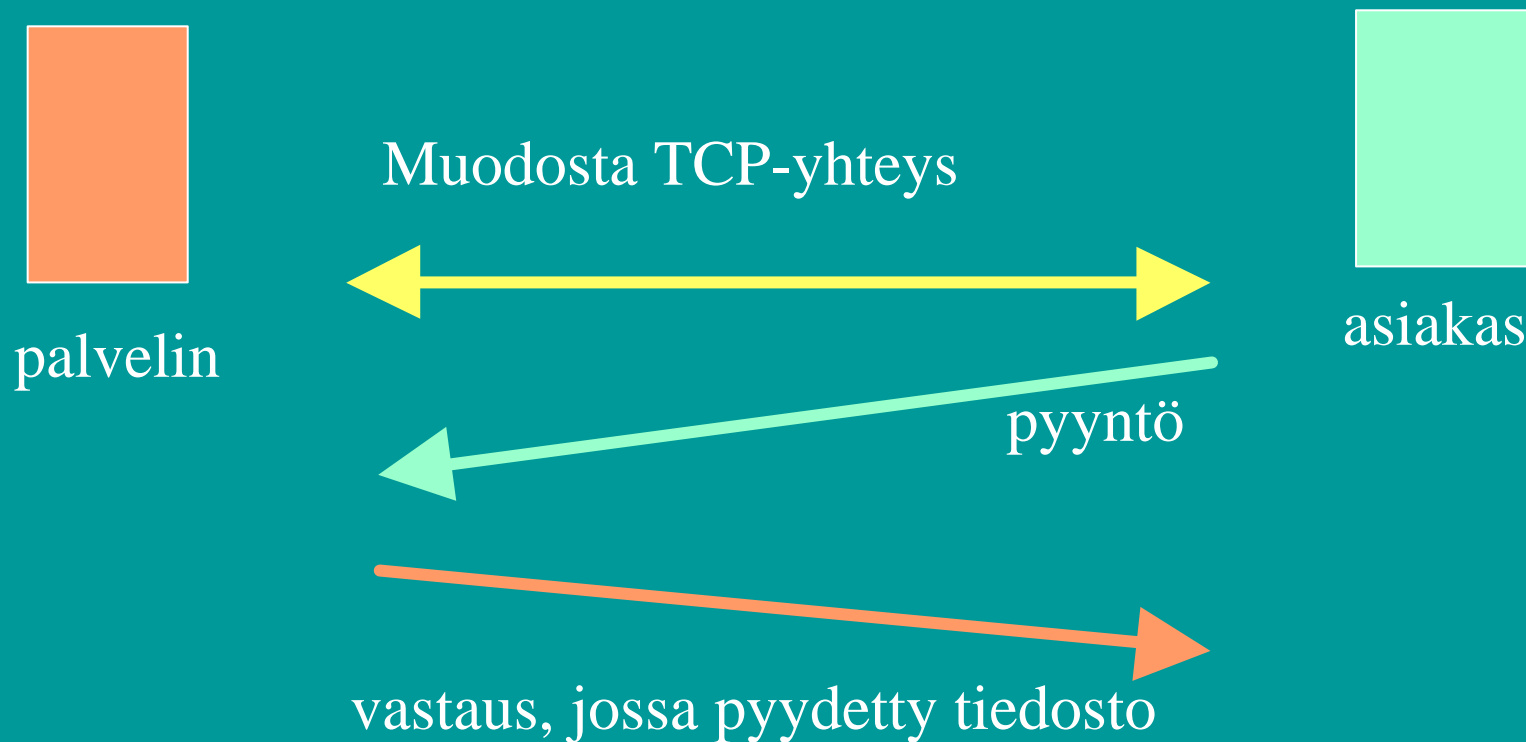
2.1. WWW

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

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

Muodostetaan TCP- yhteys koneeseen:
www.cs.princeton.edu

Pyydetään sieltä tiedosto käyttäen HTTP-protokollaa:
[index.html](http://www.cs.princeton.edu/index.html)



HTTP (Hypertext Transfer Protocol)

- **Protokolla asiakkaan ja palvelimen väliseen kommunikointiin**
 - www-sivujen hakemiseen
 - HTML-tiedosto, jossa
 - HTML-tiedostoja
 - JPEG-kuva, GIF-kuva
 - Java-sovelma
 - sivuilla URL-osoite
 - **tilaton protokolla**

23.3.2001



URL (Uniform Resource Locator)

- [Http://www.nmib.com/glossary.index.html](http://www.nmib.com/glossary.index.html)

↑
verkko-palvelin
jos puuttuu, niin
selain lisää
automaattisesti

↑
**verkkopaikan
nimi**

↑
**Organisaation
nimi = domain-
osoite**

↑
hakemisto

↑
tiedosto

Polkunimi haluttuun dokumenttiin

•
•
•

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)

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
 - vastaanottaa palvelimen lähettämän vastaussanoman, 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**

•
•
•

Palvelinprosessi

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

•
•
•

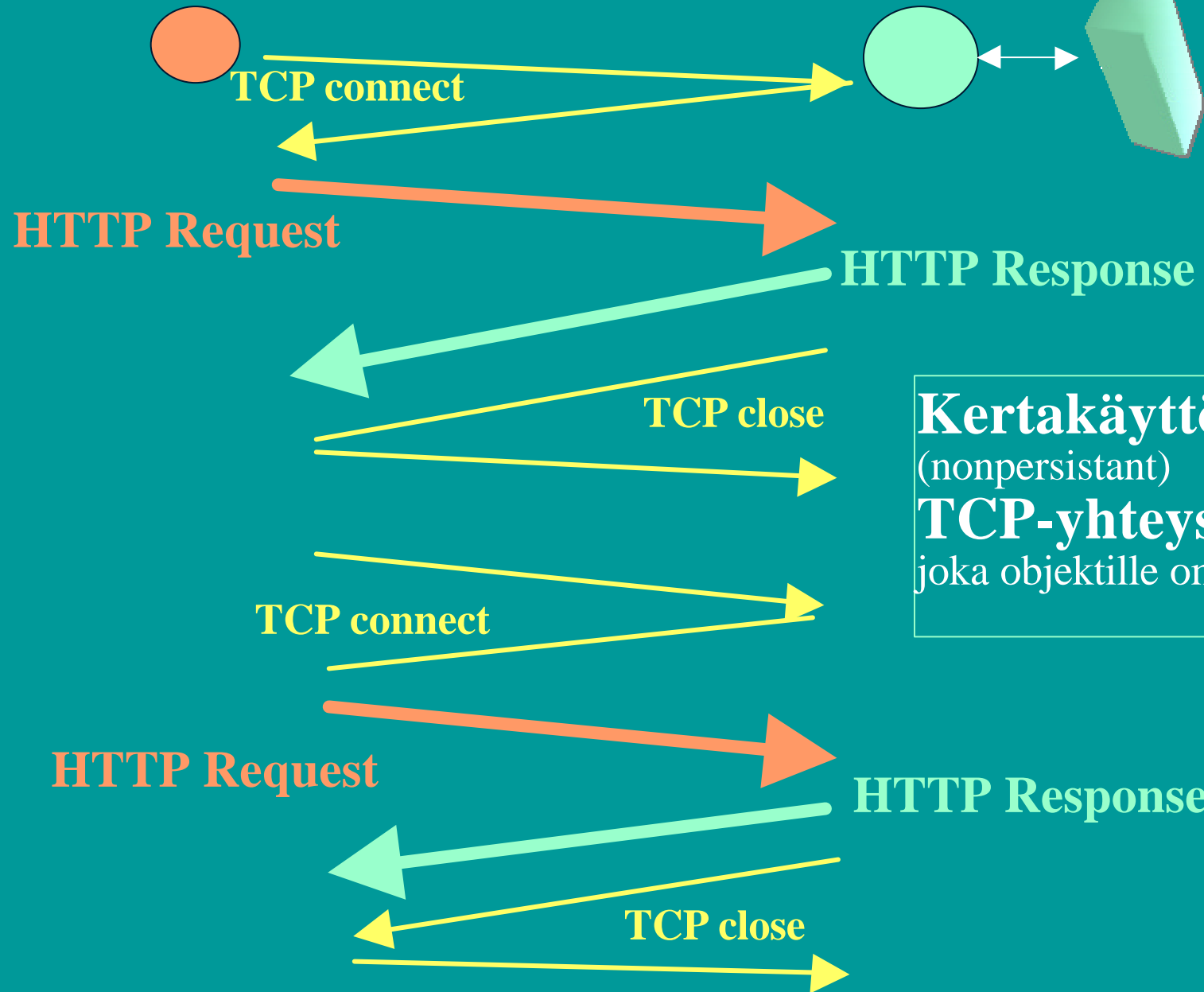
HTTP-pyyynnön tullessa

- **verkkopalvelin**

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

asiakas

palvelin

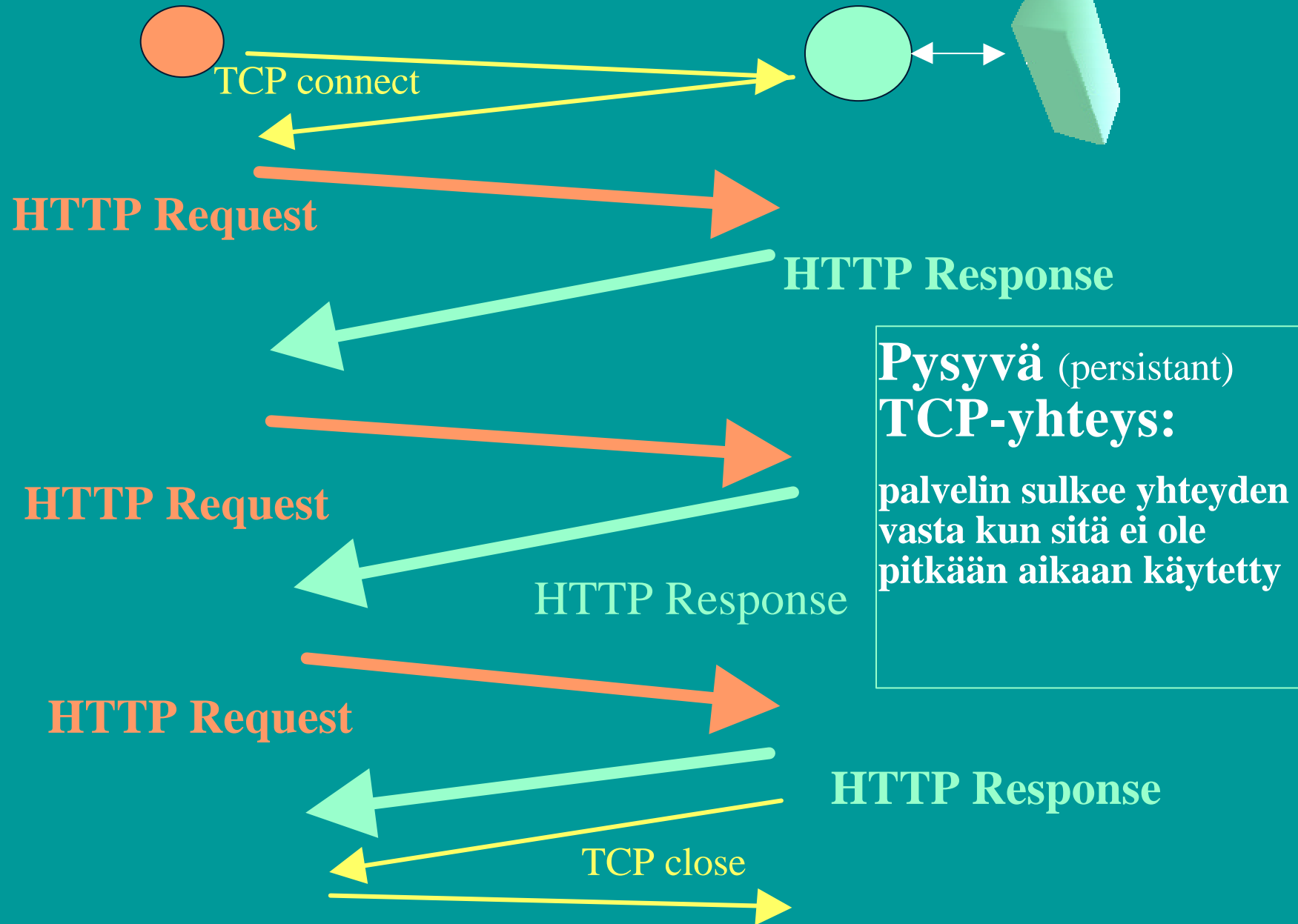


Suorituskyky?

- **Jos haetaan 10 objektia**
 - 10 TCP-yhteyden muodostusta
 - kukin 2 sanomaa = RTT => 10 RTT
 - hidas aloitus hidastaa lähetysnopeutta
 - voidaan avata useita rinnakkaisia yhteyksiä
 - puskuritilat yhteyksille
- **käytetään pysyvää TCP-yhteyttä**
 - palvelin jättää yhteyden sulkematta
 - muut pyynnöt ja vastaukset käyttävät samaa yhteyttä

asiakas

palvelin

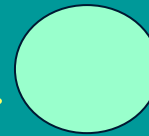


Pysyvä (persistent)
TCP-yhteys:

palvelin sulkee yhteyden
vasta kun sitä ei ole
pitkään aikaan käytetty

asiakas

palvelin



TCP connect

HTTP Request

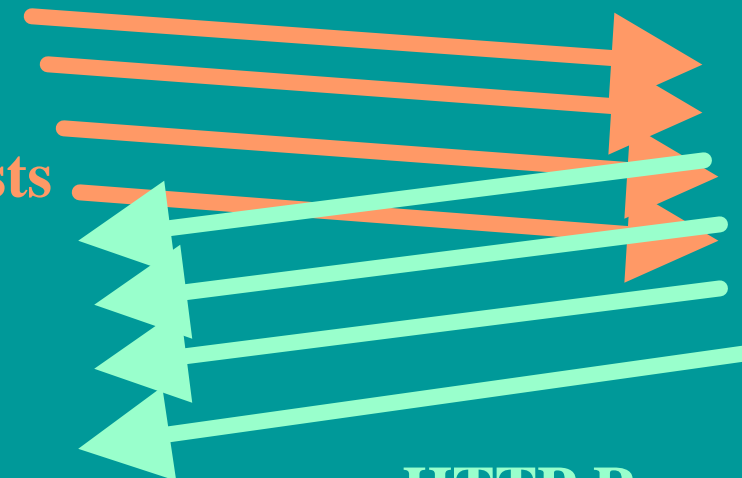
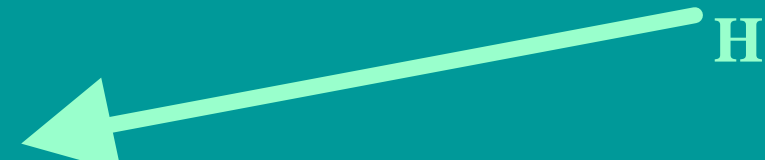
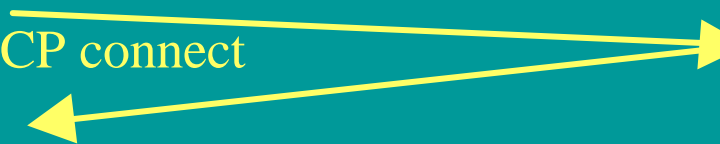
HTTP Response

HTTP Requests

Pysyvä (persistent)
TCP-yhteys
liukuhihnana (with
pipelining):
HTTP/1.1

HTTP Responses

TCP close



•
•
•

HTTP-sanomat (esimerkki)

- **HTTP Request Message:**

GET /jokuhakemisto/sivu.html HTTP/1.1

Host: WWW.jokupaikka.fi

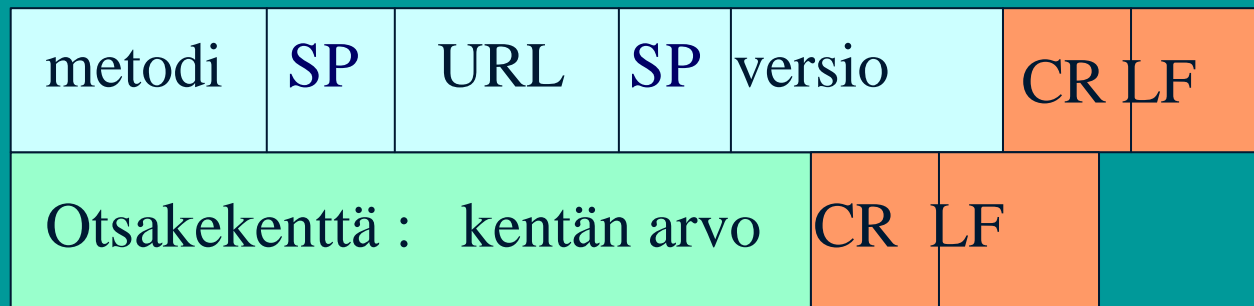
Connection: close

User-agent: Mozilla/4.0

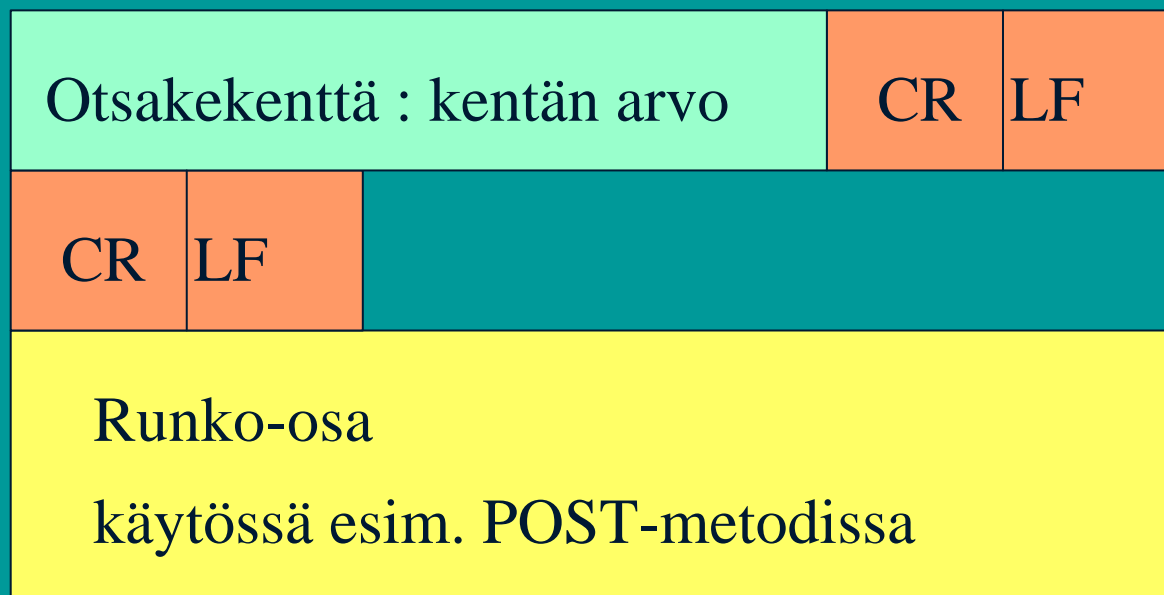
Accept-language:fi

(ylimääräinen CR ja LF)

Pyyntösanomien yleinen rakenne



... Lisää otsakerivejä



•
•
•

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)

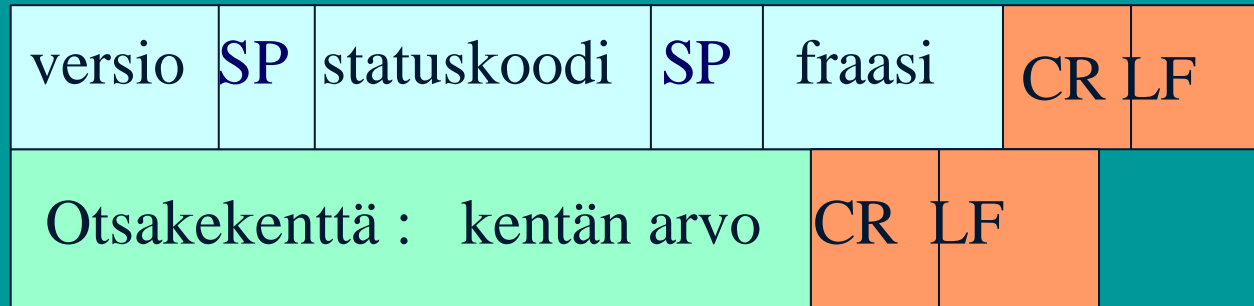
Last-Modified: Mon, 22 Jun 2000 09:23:24 GMT

Content-Length: 6821

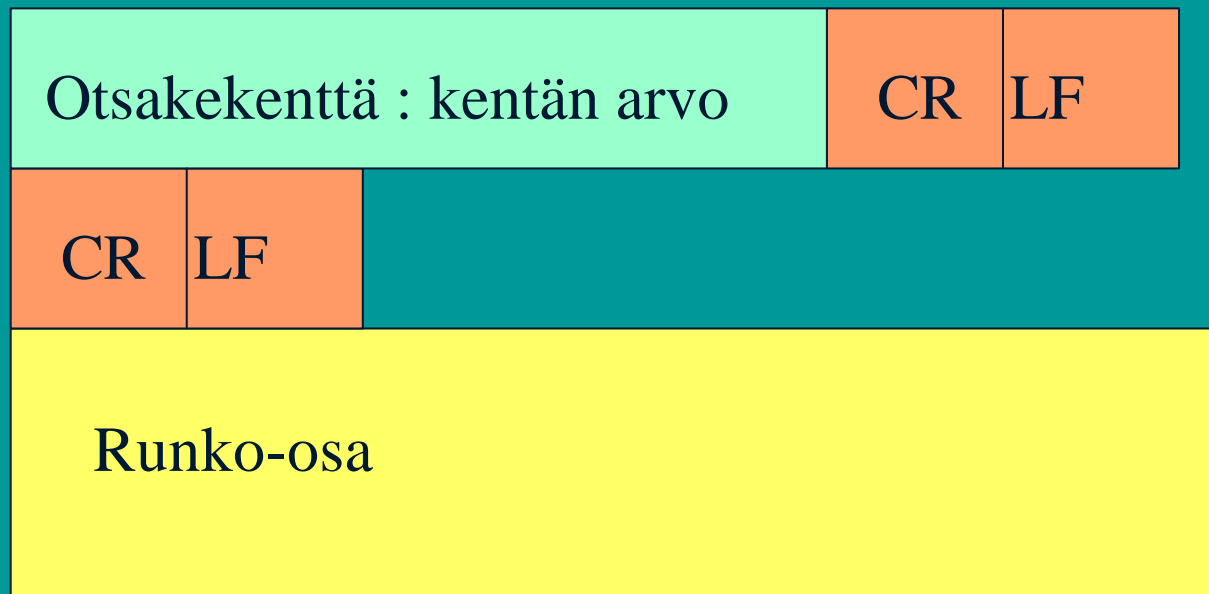
Content-Type: text/html

(data data data data data ...)

Vastaussanomman yleinen rakenne



... Lisää otsakerivejä



•
•
•

Pyyntömetodeja

- **GET** sivun lukeminen
 - aina tai vain tietyn päiväyksen jälkeen muutettu sivu
- **HEAD** pelkän otsikon lukeminen
- **PUT** sivun tallettaminen
- **POST** lisäys verkkosivulle
- **DELETE** sivun poistaminen
- **LINK** uusi linkki
- **UNLINK** linkin poistaminen

•
•
•

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**

Käyttäjän tunnistaminen

- **Autentikointi**

- omat koodit ja otsakkeet tätä varten

- 401 Authorization Required -statuskoodi

WWW-Authenticate -otsake kertoo kuinka

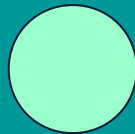
- asiakas lähettää autentikointitiedot joka pyynnössä (Authorization- otsakerivi)

- **pipari** (cookie)



- asiakkaalle talteen käyttäjäkohtainen tunnuskoodi, joka esitetään joka pyynnön yhteydessä

asiakas



...

Set-cookie: 167845<CR><LF>

....



Pipari-
tiedosto



...

Cookie: 167845<CR><LF>

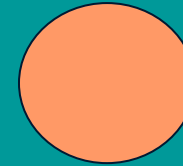


...

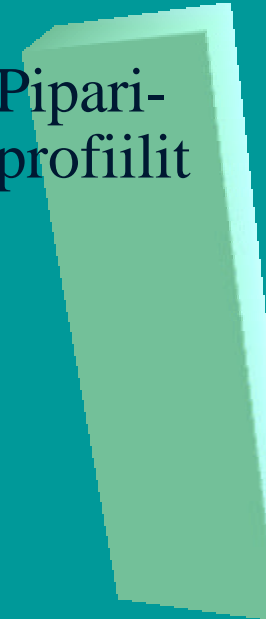
Cookie: 167845<CR><LF>



palvelin



Pipari-
profiilit



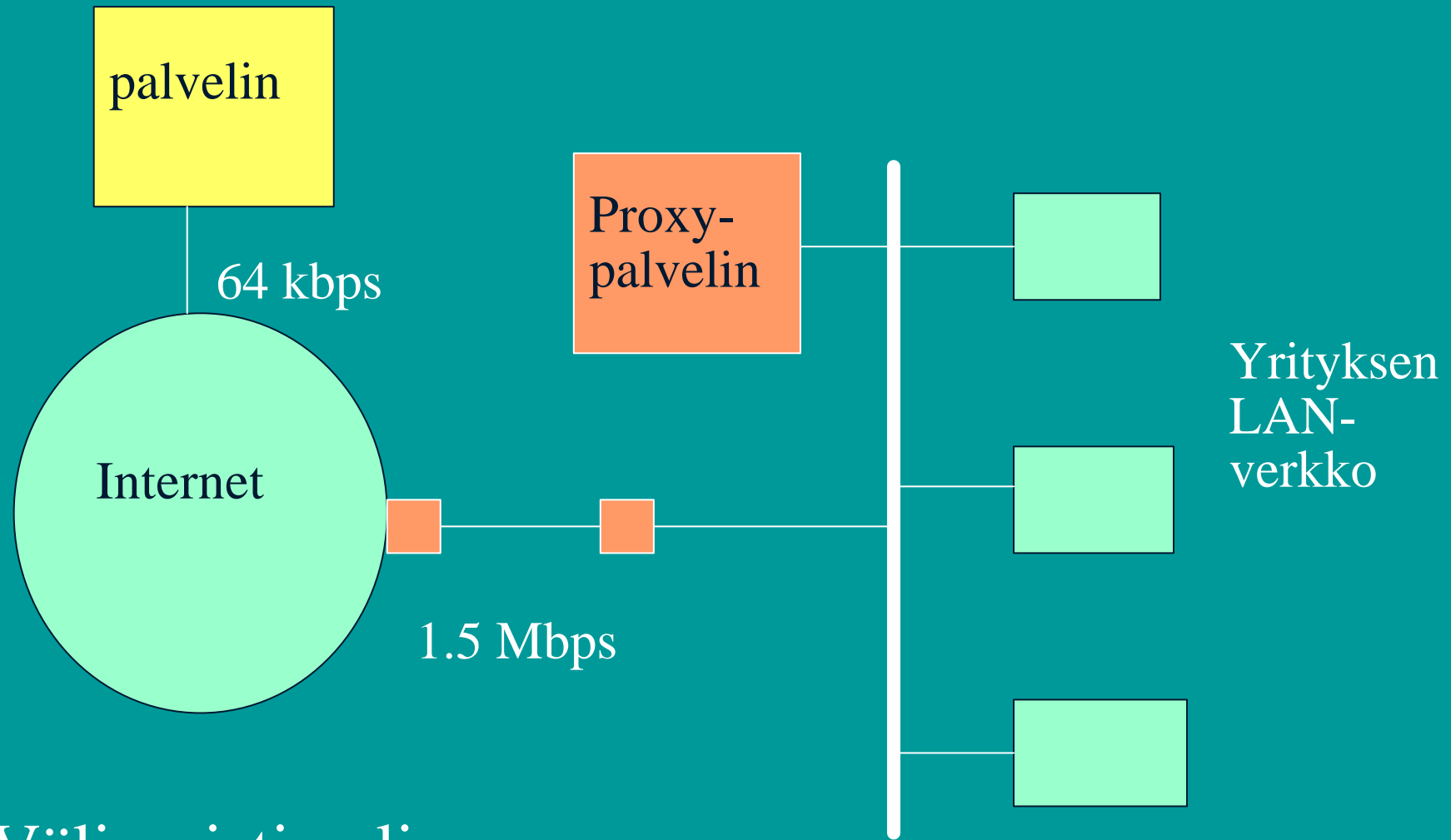
Käyttö:

Autentikointi,
kohdistettu mainonta,
virtuaalinen ostoskärry

•
•
•

Verkkovälimuisti

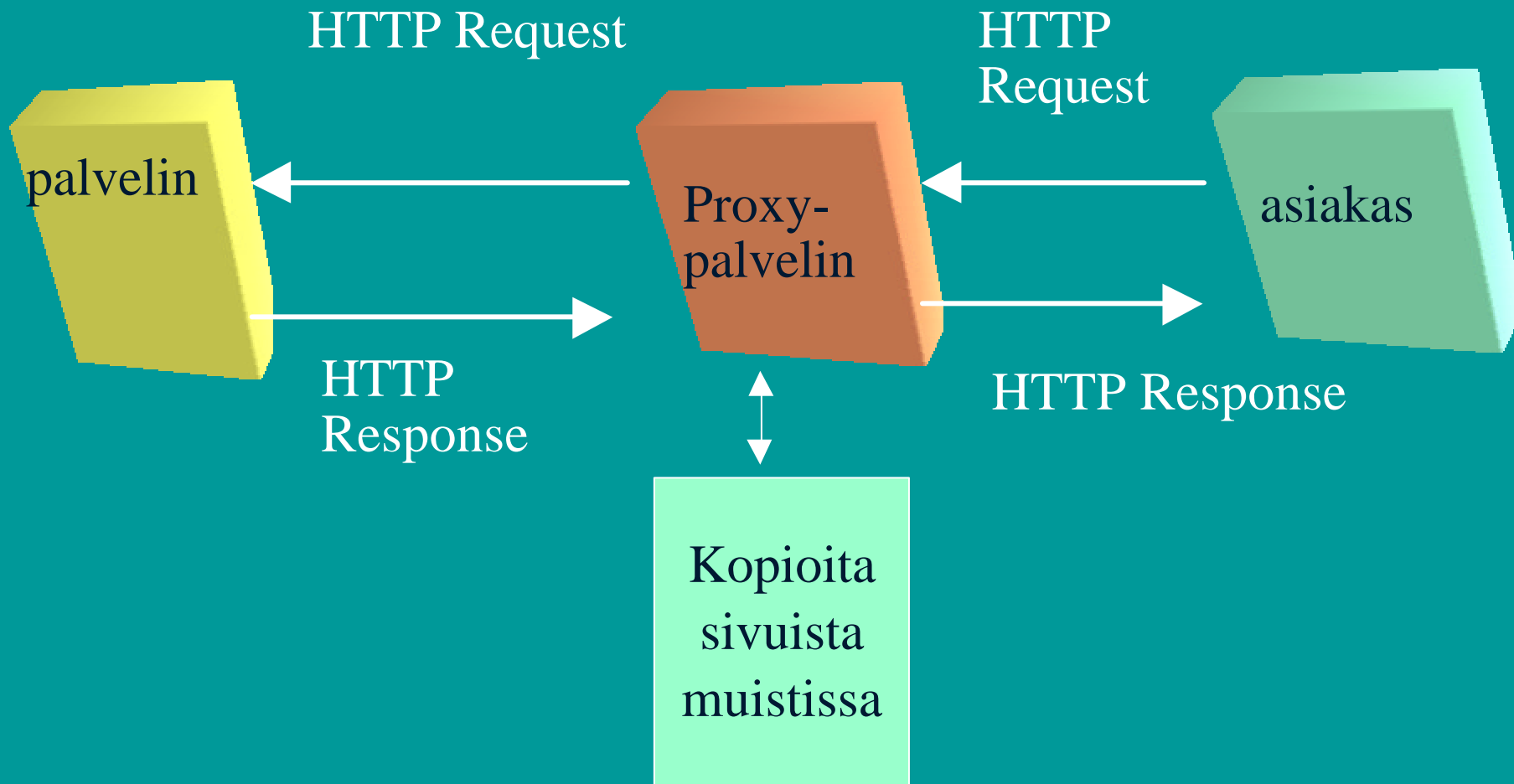
- **Säilyttää kopioita haetuista sivuista**
 - viimeksi haetut sivut
- **pyyntö ohjautuu ensin välimuistiin**
 - haetaan verkon yli vasta, jos ei löydy välimuistista
- **etuja:**
 - lyhentää kyselyaikaa
 - vähentää Internetin liikennettä
 - tuo nopean jakelukanavan hitaille palvelimille



Välimuistin eli ns.
Proxy-palvelimen
käyttö

10-100 Mbps

Haku palvelimelta vain, jos ei ole jo muistissa



-
-
-

Conditional GET

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

-
-
-

HTML (HyperText Markup Language)

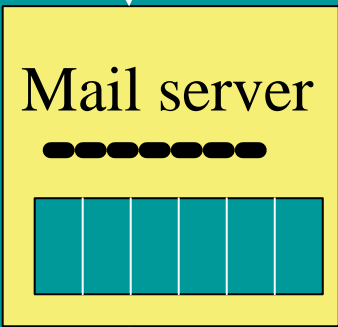
- **SGML (Standard Generalized Markup Language)**
 - merkkäuskieli
 - kertoo, kuinka dokumentit muotoillaan
 - TeX, troff,
 - ladontamerkinnot
 - WYSIWYG
- **XML (Extensible Markup Language)**

•
•
•

2.2. Sähköposti

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

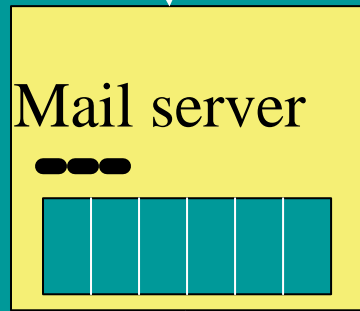
User agent



User agent

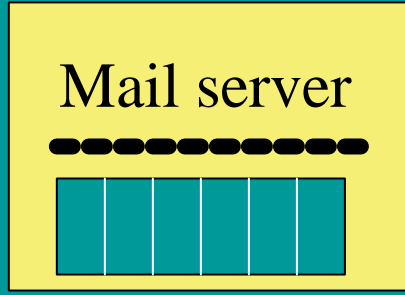
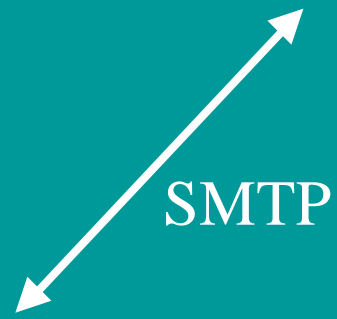
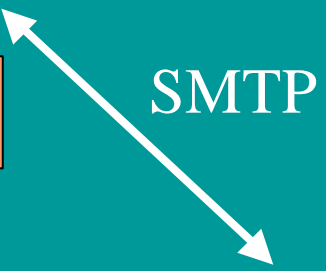


User agent



User agent

User agent



User agent

.....

Lähtevien sanomien jono



SMTP (Simple Mail Transfer Protocol)

- **Postipalvelimet käyttävät SMTP-protokollaa sanomien siirtoon**
 - sanoma lähetävä postipalvelu muodostaa TCP-yhteyden vastaanottajan postipalvelimeen
 - ‘kättelyjen’ jälkeen siirretään sanomat

Vastaan-
ottava on
palvelin



Lähetävä on
asiakas

Lähettävä postipalvelin

- muodostaa TCP-yhteyden
- esittelee itsensä
- kertoo lähettäjän ja vastaanottajan
- lähettää sanoman, jos vastaanottaja antaa luvalla

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 <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

-
-
-

- **TCP-yhteys on pysyvä**

- yhteydellä voidaan lähettää useita sanomia

- **Yhteydellä voidaan lähettää vain ASCII-sanomia**

- sanomatekstin lopettaa omalla rivillään irrallisena oleva piste (“.”)

- “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

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

-
-
-

Sanoman muoto

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

Yleisiä sanoman
otsakekenttiä

kentät erotettu
rivinvaihdolla

RFC 822 sallii myös omien
kenttien käytön

-
-
-

MIME (Multipurpose Internet Mail Extension)

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

-
-
-

MIME-laajennus

- **Lisää kenttiä otsikkoon**

MIME-versio:

Content-Transfer-Encoding:

Content-Type:

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)

▪

-
-
-

MIME-sisältötyyppejä

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

Content-Type: multipart/mixed;Boundary=StartOfNextPart

-- StartOfNextPart

Hei Allu,
sinulle kaunis kuva kissastani.

-- StartOfNextPart

Content-Transfer-Encoding: base64

Content-Type: image/jpeg

base64 encoded data

.....

.....base64 encoded data

-- StartOfNextPart

Haluatko muita kuvia!

.

-
-
-

- **tyyppien koodaukset siirtoa varten**
 - sähköpostisanomassa saa olla vain ASCII-merkkejä
 - ei esimerkiksi ääkkösiä
 - **base64-koodaus**

•
•
•

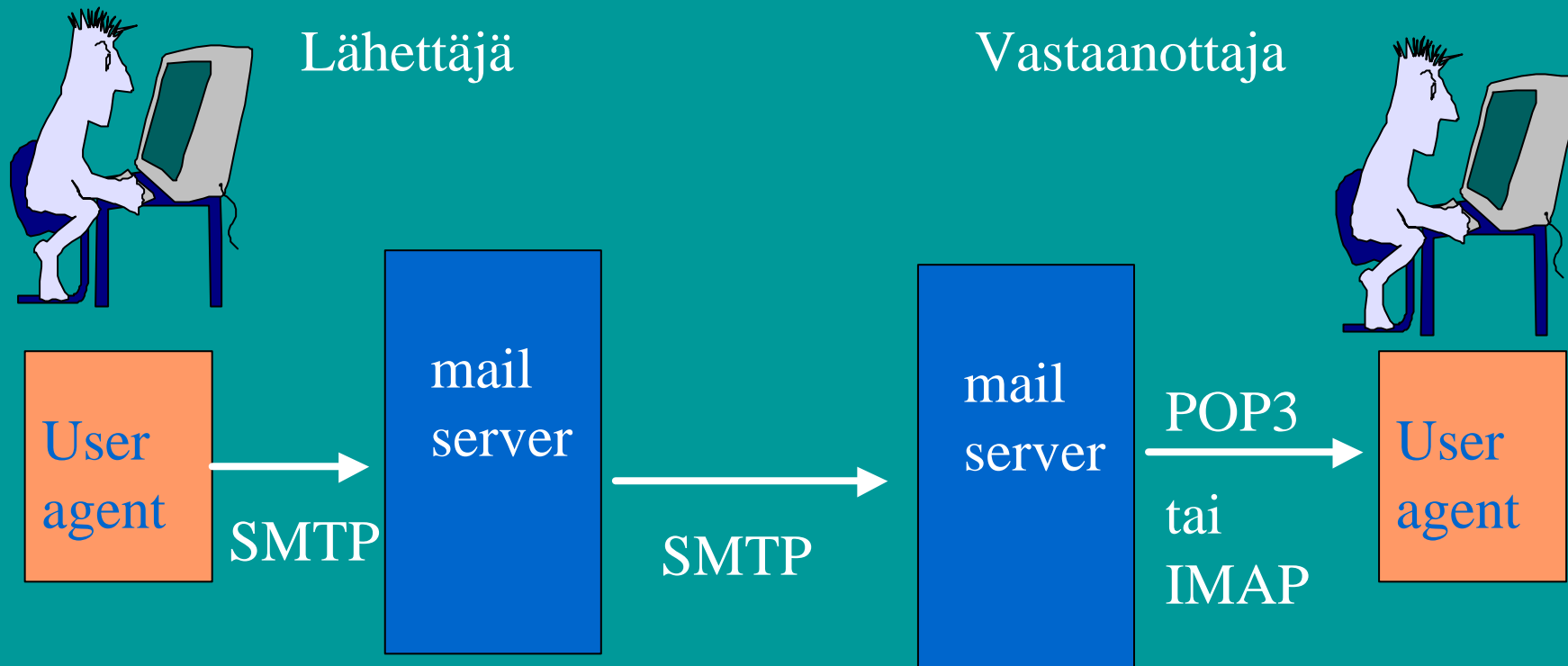
Base64-koodaus

- **Sanoman 24 bitin ryhmät jaetaan 6 bitin osiksi, jotka kukin koodataan ASCII-merkiksi => 64 eri vaihtoehtoa**
 - 0-25 = A-Z,
 - 26-51 = a-z,
 - 52-61 = 0-9,
 - 62= +, 63 =/
- Tekstisanoma voidaan koodata myös 7-bitin ASCIIlla.

•
•
•

Postiprotokollat (mail access protocols)

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



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

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

•
•
•

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.marttinen@helsinki.fi
 - vastaanottajan kone usein pois päältä

•
•
•

2.3. DNS (Domain Name System)

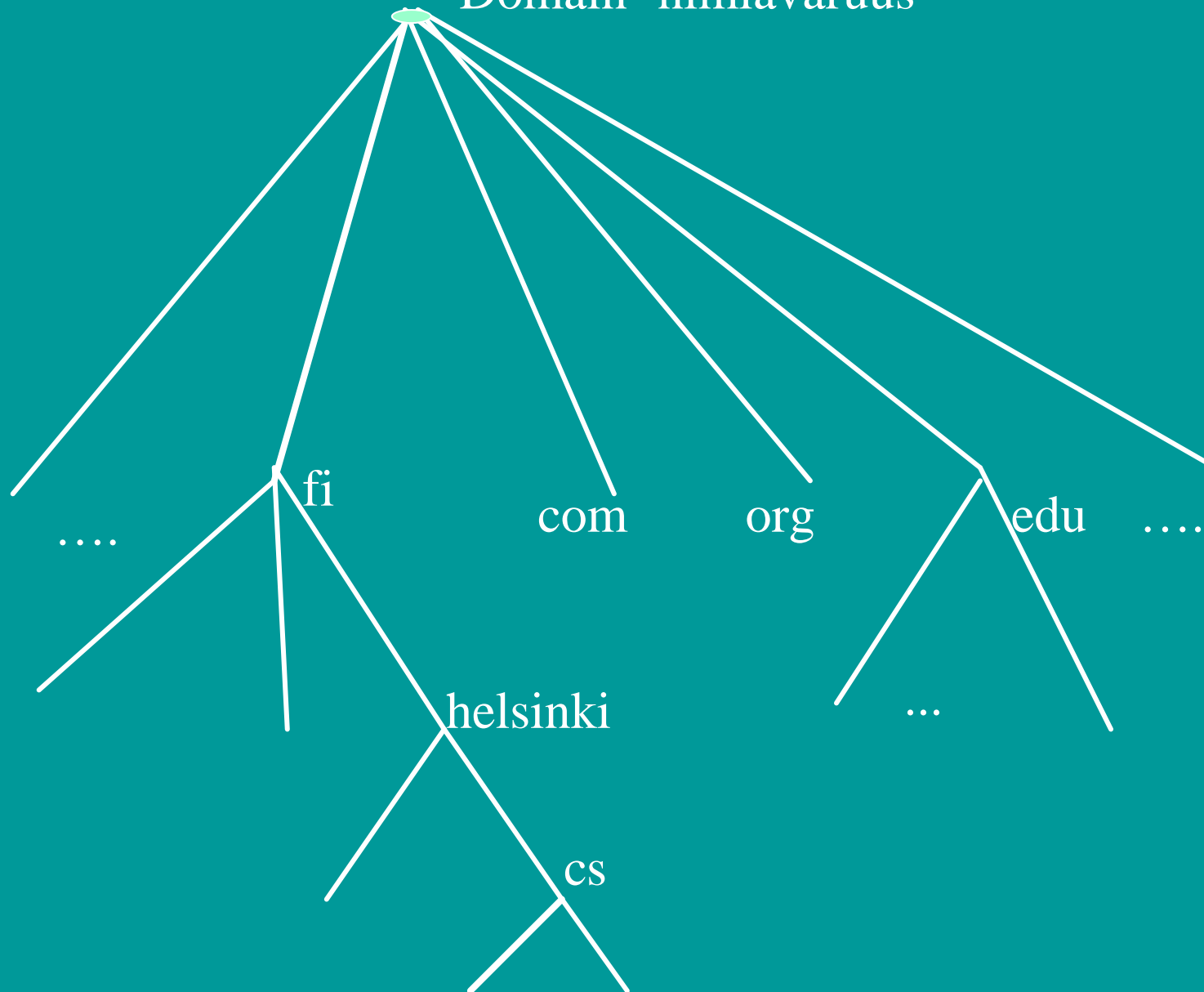
- **Internetin hakemistopalvelu**
- Ihmisten ja ohjelmien käyttämien merkkijonomuotoisten nimien muuttaminen binäärisiksi IP-verkko-osoitteiksi
 - **hydra.cs.helsinki.fi => 218.214.4.29**
- hierarkkinen nimentätapa + hajautettu hakemisto IP-osoitteiden selvittämiseksi
 - **esim. sähköpostiosoitteita,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

-
-
-

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.u

Domain -nimiavaruus



•
•
•

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

•
•
•

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

-
-
-

- **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>,

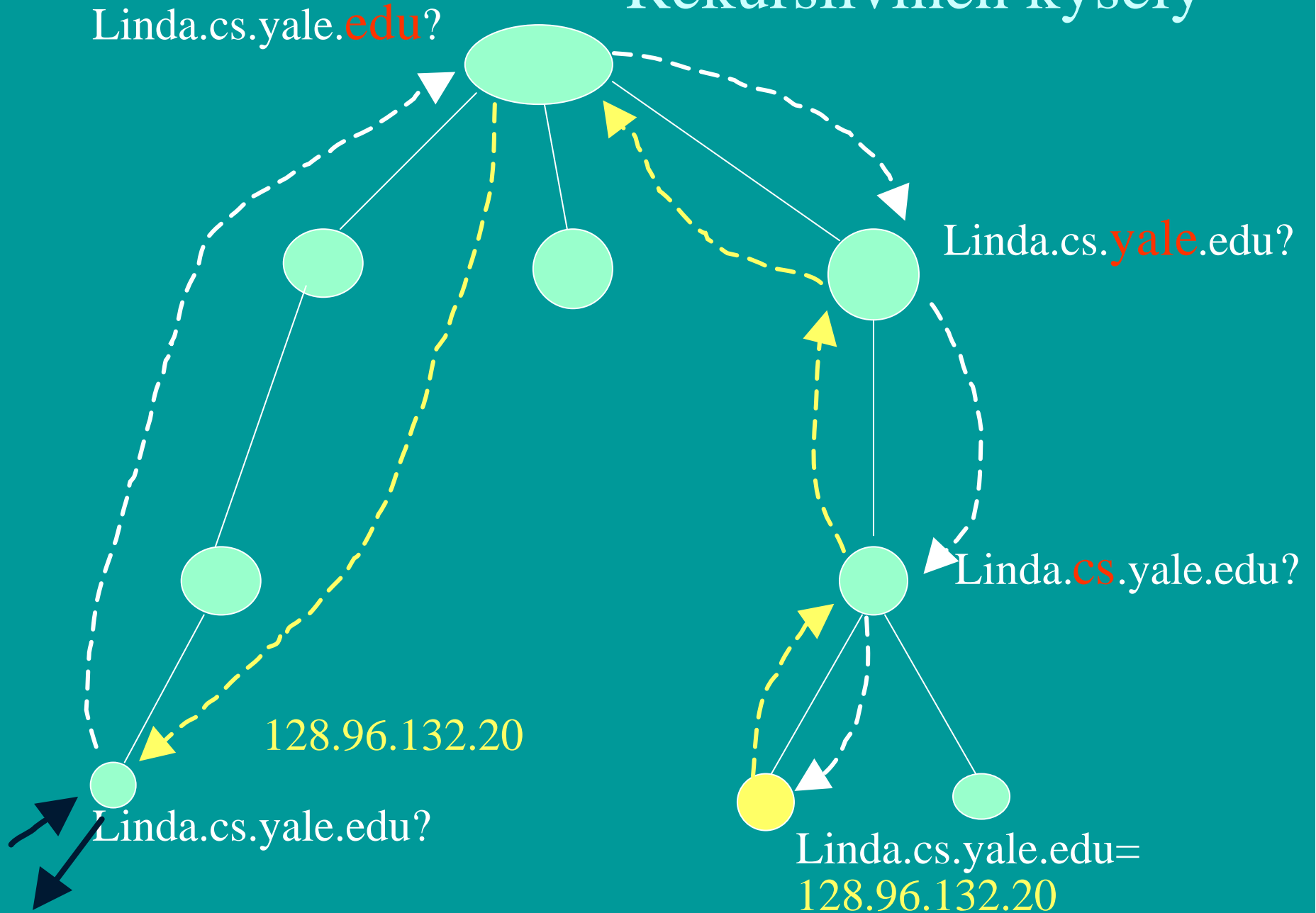
<ftp://ftp.rs.internic.net/domain/named.root>)

- esim. Kysely “linda.cs.yale.edu” lähetetään nimipalvelimelle “edu-server.net”, joka vastaa vyöhykkeestä “edu”.

-
-
-

- **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

Rekursiivinen kysely



Iteratiivinen kysely

Linda.cs.yale.edu?
Kysy: 128.99.138.70:lta

Linda.cs.yale.edu?

Kysy:
128.108.37.190:lta

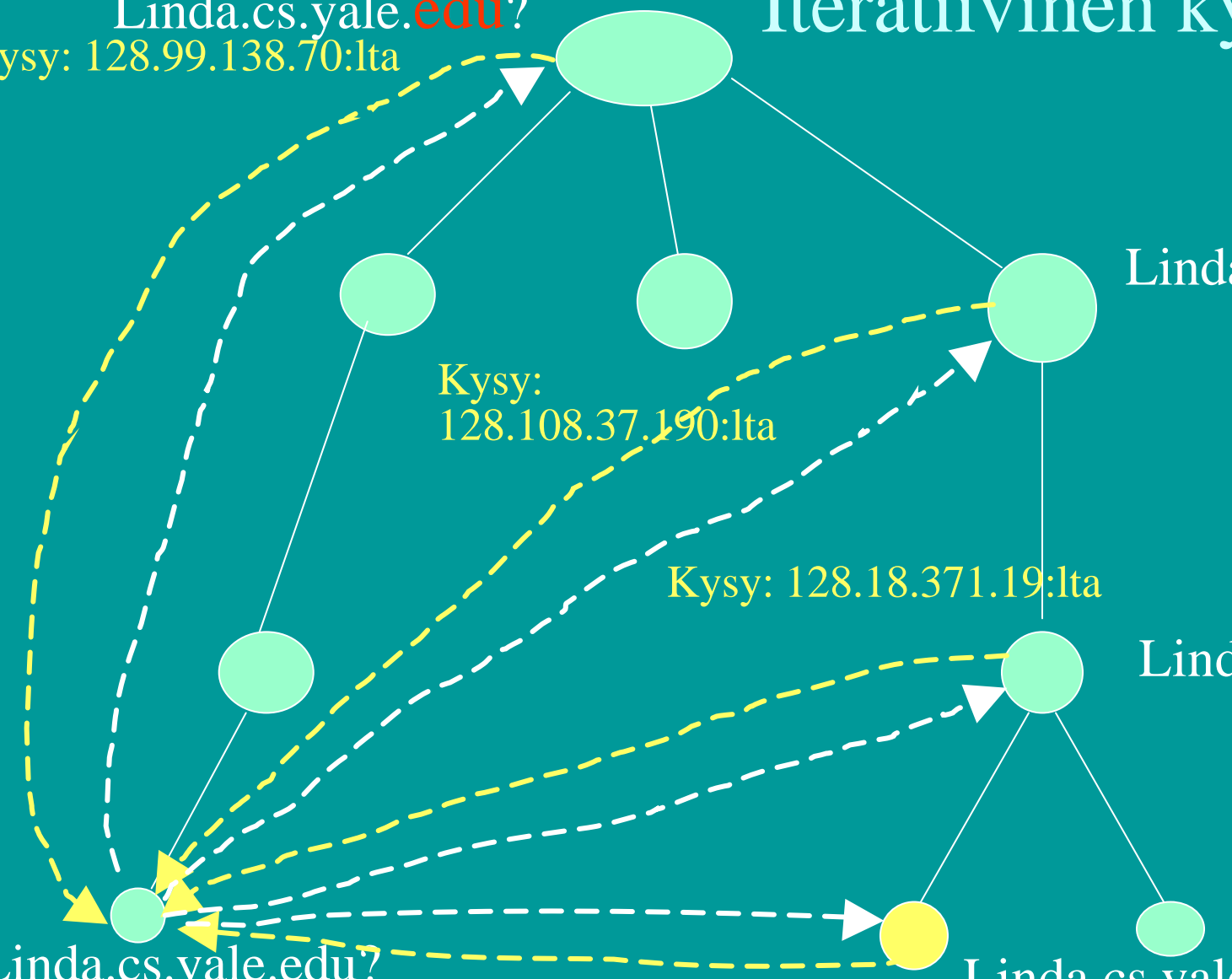
Kysy: 128.18.371.19:lta

Linda.cs.yale.edu?

Linda.cs.yale.edu?

128.96.132.20

Linda.cs.yale.edu=
128.96.132.20



•
•
•

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

-
-
-

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	

•
•
•

DNS-välimuisti

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