



# Tietoliikenteen perusteet

## Kuljetuskerros

Kurose, Ross: Ch 3



# Sisältöä

- n Kuljetuspalvelut
- n Yhteydetön kuljetuspalvelu, UDP
- n Luotettavan kuljetuspalvelun periaatteet
- n Yhteydellinen kuljetuspalvelu, TCP
- n Ruuhkanhallinta TCP:ssä

## Oppimistavoitteet:

- Tuntea Internetin kuljetusprotokollien (UDP/TCP) toiminnallisuus ja periaatteet
- Osata luotettavan kuljetuspalvelun ja vuonvalvonnan periaatteet ja toteutukset
- Osata TCP-ruuhkanhallinnan





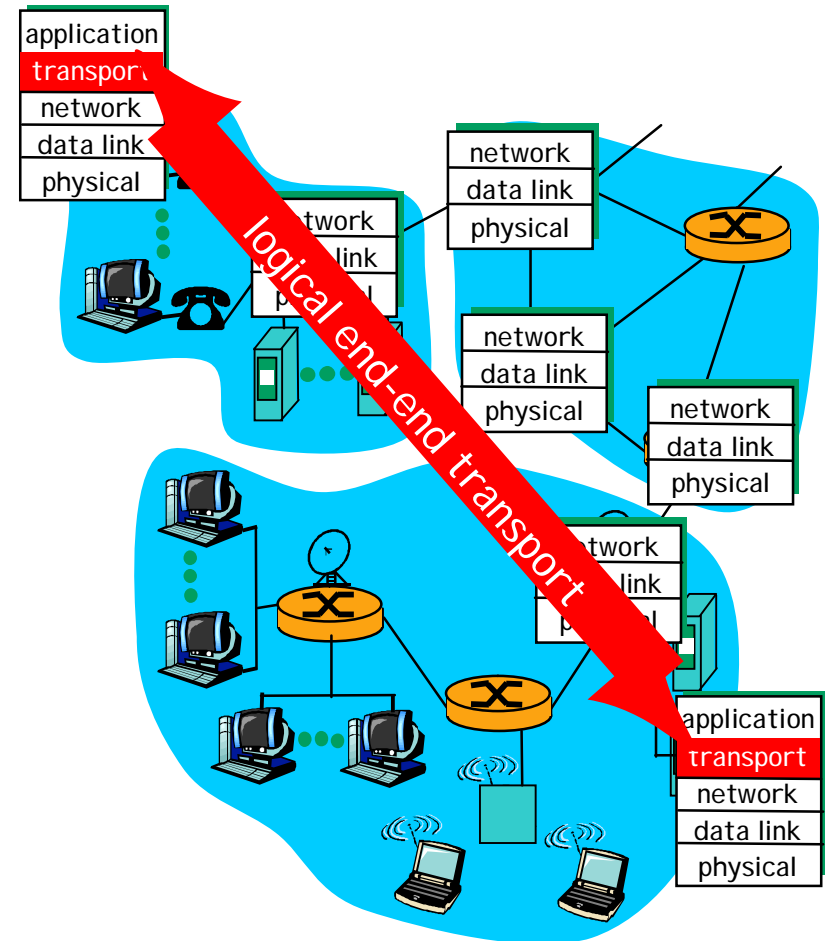
# Kuljetuskerros

## Kuljetuspalvelu



# Kuljetuskerros

- n Tarjoaa kuljetuspalvelun prosessien välille
- n Vain isäntäkoneissa  
Lähetys: Pilko sovelluskerroksen sanoma pienemmiksi segmenteiksi, jotka verkkokerros toimittaa perille.  
Vastaanotto: Kokoa segmentit sanomaksi, jonka sovellus lukee.
- n Verkkokerros reitittää koneesta koneelle
- n Segmentin koko s.e. verkkokerros pystyisi välittämään sellaisena



KuRo08: Fig 3.1



# Sovelluksen vaatimuksia

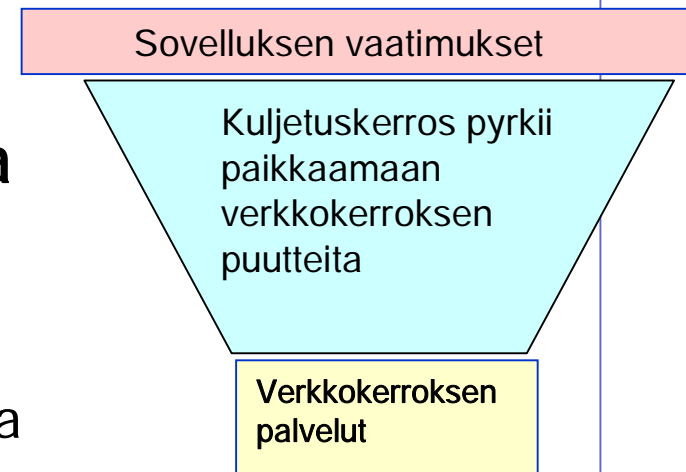
## n Verkkokerroksen palvelu voi

- n Muuttaa segmentin bittejä tai kadottaa segmenttejä
- n Toimittaa segmentit epäjärjestyksessä kuljetuskerrokselle
- n Viivyyttää segmenttejä satunnaisen pitkän ajan
- n Luovuttaa kuljetuskerrokselle useita kopioita samasta segmentistä
- n Rajoittaa segmentin kokoa

## n Sovellus edellyttää kuljetuspalvelulta

- n Virheettömyyttä, luotettavuutta
- n Järjestyksen säilymistä
- n Kaksoiskappaleiden karsimista
- n Mielivaltaisen pitkien sanomien sallimista
- n Vuonvalvonnan mahdollistamista

## n Kuljetuskerros peittää verkkokerroksen puutteita ja parantaa sovelluksen näkemää palvelun laatua



# Internetin kuljetusprotokollat

## n TCP: luotettava, järjestyksen säilyttävä tavujen kuljetuspalvelu

- n Virheenvalvonta (error control): Huomaa ja korjaa virheet, hylkää kaksoiskappaleet
- n Vuonvalvonta (flow control): Älä ylikuormita vastaanottajaa
- n Ruuhkanhallinta (congestion control): Älä ylikuormita verkkoa
- n Yhteyden muodostaminen ja purku

## n UDP: Ei-luotettava, ei-järjestyksen säilyttävä sanomien kuljetuspalvelu

- n Välittää vain sanomia, ei pyri mitenkään parantamaan verkkokerroksen tarjoamaa palvelun laatua
- n Luotettavuus jää sovelluskerroksen hoidettavaksi

## n Kumpikaan kuljetuspalvelu ei anna takuita viiveelle tai siirtonopeudelle ("best effort")



# Mikä kone /Mikä prosessi?

- n Kuljetuskerros tarjoaa päästä-päähän yhteyden
  - n Prosessilta prosessille ( = pistokkeesta pistokkeeseen)
  - n Prosessi lukee ja kirjoittaa sanomia halutessaan
- n Datan lisäksi on välitettävä osoitetietoja
  - n Vastaanottajan ja lähettäjän tiedot
  - n Eri koneiden prosessit voivat käyttää samaa palvelua
  - n Saman koneen prosessit voivat käyttää eri palveluita
- n Kuljetuskerros: mikä prosessi = mikä portti
- n Verkkokerros: mikä kone = mikä IP-osoite
- n Porttinumero
  - n 16-bittinen: 0 – 65535
  - n Portit 0 – 1024 on varattu kukin tietyille palvelulle (well known ports)
    - Esim. www-palvelulle portti 80, SMTP-postipalvelulle portti 25

# Mikä kone /Mikä prosessi?

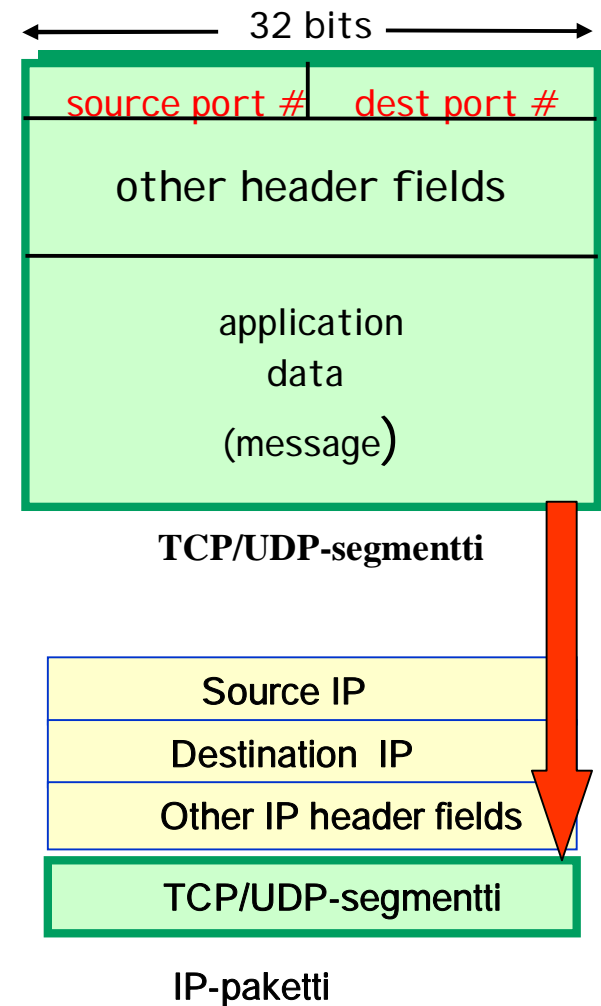
## Lähetys (asiakas)

### n Kuljetuskerros

- n Segmentin otsakkeessa lähde- ja kohdeprosessin porttinumero
- n Antaa segmentin verkkokerroksen välitettäväksi
- n TCP: huolehtii myös luotettavuudesta
- n UDP: tarjoaa pelkän välityspalvelun

### n Verkkokerros

- n Paketin otsakkeessa lähde- ja kohdekoneen IP-osoite → reitittimet osaavat ohjata oikealle koneelle





# ▪ Mikä kone /Mikä prosessi?

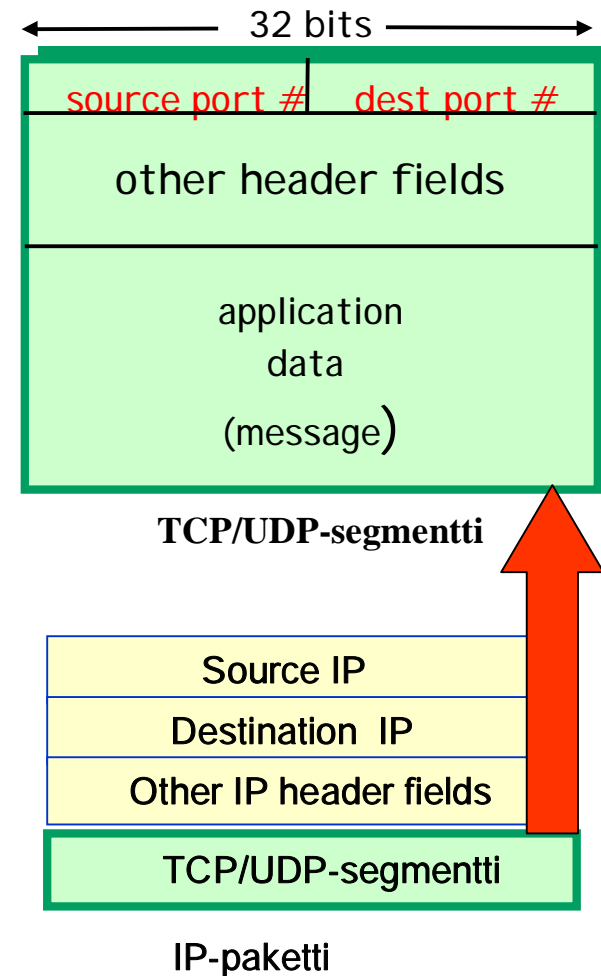
## Vastaanotto (palvelija)

### n Verkkokerros

- n Vastaanottaa IP-paketin
- n Poistaa verkkokerroksen otsaketiedot
- n Luovuttaa paketissa olleen segmentin kuljetuskerrokselle

### n Kuljetuskerros

- n Poistaa kuljetuskerroksen otsaketiedot
- n Kokoaa yhteenkuuluvat segmentit sanomiksi (tavuvirraksi)
- n Ohjaa sanoman (tavuvirran) oikealle prosessille (eli oikeaan pistokkeseen) porttinumeron avulla
  - TCP: huolehtii myös luotettavuudesta
  - UDP: tarjoaa pelkän välityspalvelun





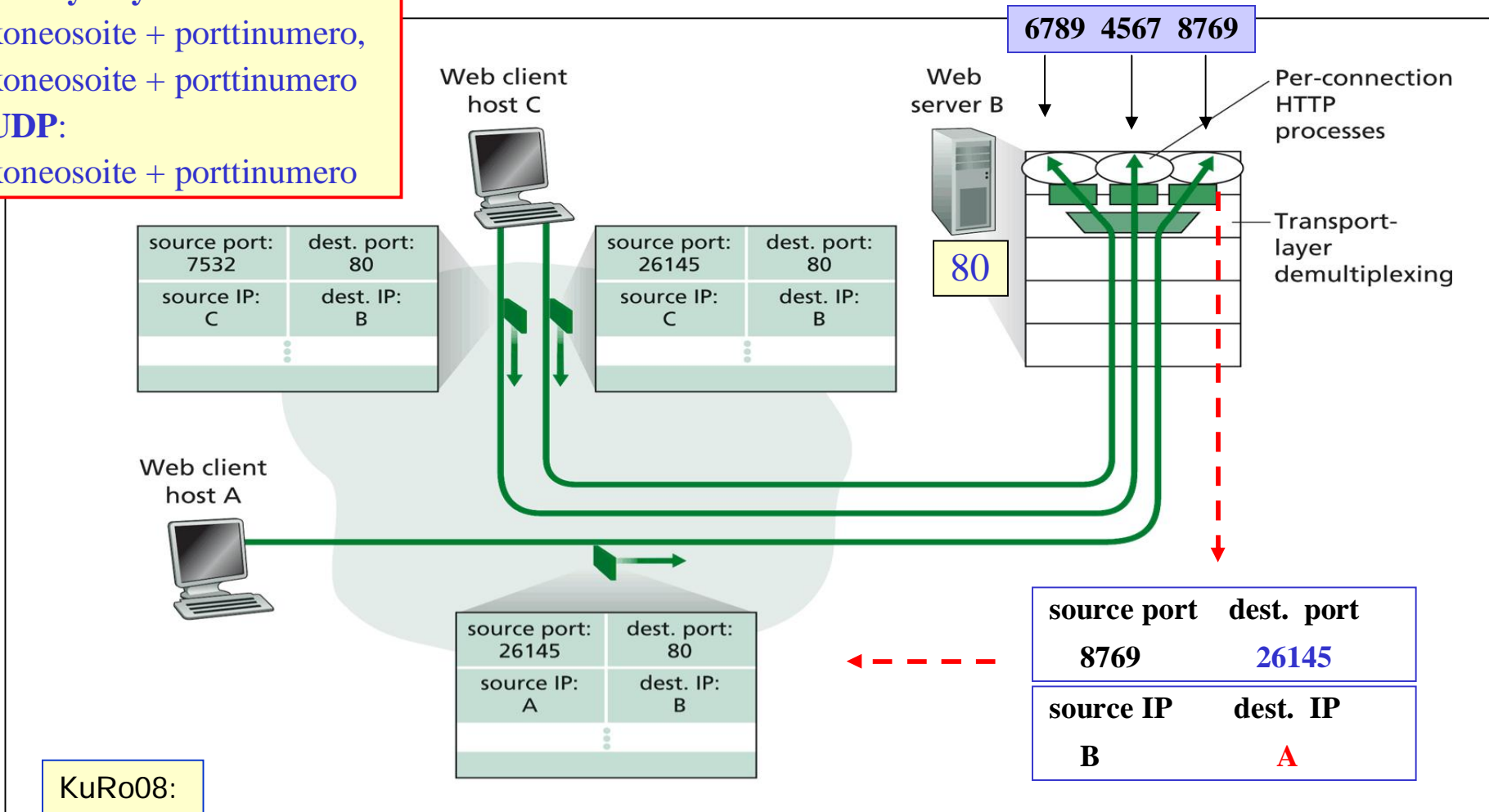
# Kaksi www-asiakasta ja palvelija

## TCP-yhteys:

koneosoite + porttinumero,  
koneosoite + porttinumero

## UDP:

koneosoite + porttinumero



KuRo08:

**Figure 3.5** ♦ Two clients, using the same destination port number (80) to communicate with the same Web server application



## Kuljetuskerros

Yhteydetön  
kuljetuspalvelu  
UDP



# UDP (User Datagram Protocol) (RFC 768)

## n Yhteydetön

KJ ei pidä tallessa mitään sovellusten väliseen keskusteluun liittyvää.

Sovellus antaa aina sanoman lisäksi kohdeosoitteen ja kohdeportin

## n Ei takaa luotettavuutta (~'epäluotettava'?)

n vain minimi palvelu: mille koneelle, mihin porttiin

n voi kadottaa sanoman

## n Ei myöskään säilytä sanomien järjestystä

n Sovellus saa sanomat siinä järjestyksessä kuin ne tulevat perille

## n Vähän yleisrasitetta

n Aikaa ei kulu yhteyden muodostukseen eikä purkuun

n Ei kulu resursseja yhteyden tilatietojen ylläpitoon

n Pieni otsake (eli vähän itse protokollaan liittyviä tietoja)

n Ruuhkanhallinta ei säännöstele liikennettä



# Käyttö

**n Vaikka UDP ei takaa luotettavuutta, se sopii silti monen sovelluksen tarpeisiin**

Alkujaan oli vain TCP, UDP luotu myöhemmin

**n Miksi UDP?**

**n** Pieni yleisrasite, koska pieni otsake

**n** Ei tilatietojen tallettamista eikä lähetys- ja vastaanottopuskureita

**n** Sovellus voi sietää virheitä

**n** Reaaliaikavaatimuksia: data lähtee heti verkkoon, koska ei yhteydenmuodostusta eikä vuon- tai ruuhkanvalvontaa

**n Tarvittava luotettavuus on räätälöitävissä sovellukseen**

Sovellusprotokolla: oma sanomanumerointi, uudelleenlähetys



# Kuljetusprotokolla TCP

**nTCP** (Transmission Control Protocol) [RFC 793]

**Yhteydellinen palvelu** (connection-oriented)

Yhteyden muodostus ennen datan siirtoa (handshaking)

Kaksisuuntainen TCP-yhteys (full-duplex)

Yhteyden purku (shutdown)

**Luotettava kuljetuspalvelu**

Järjestyksen säilyttävä tavuvirta sovellukselle

segmenttinumerot, kuittaukset, uudelleenlähetykset

**Vuonvalvonta** (flow control)

Lähettäjä hiljentää vauhtia, jos **vastaanottaja** ei ehdi käsitellä

**Ruuhkanvalvonta** (congestion control)

Lähettäjä hiljentää vauhtia, jos **reitittimet** eivät ehdi käsitellä



# Verkkosovelluksia

Application	Application-Layer Protocol	Underlying Transport Protocol
Electronic mail	SMTP	TCP
Remote terminal access	Telnet	TCP
Web	HTTP	TCP
File transfer	FTP	TCP
Remote file server	NFS	Typically UDP
Streaming multimedia	typically proprietary	<b>UDP or TCP</b>
Internet telephony	typically proprietary	<b>UDP or TCP</b>
Network management	SNMP	Typically UDP
Routing protocol	RIP	Typically UDP
Name translation	DNS	Typically UDP

**Figure 3.6** ♦ Popular Internet applications and their underlying transport protocols

Kuro08:

Miksi nämä sovellukset suosivat UDP:tä?



# UDP-segmentin rakenne

## n Porttinumerot

- n Koska on prosessien välinen palvelu

## n Length

- n Segmentin kokonaispituus  
otsake (8 B) mukaanluettuna

## n Checksum (optionaalinen)

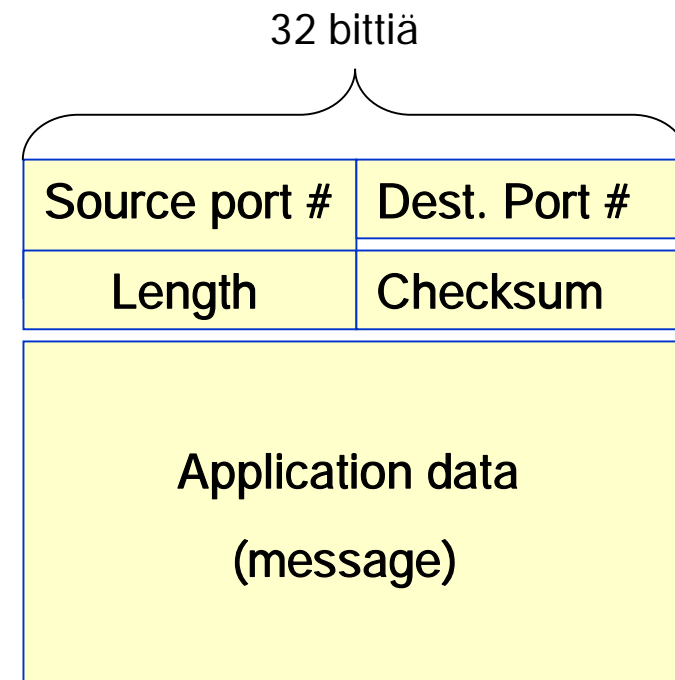
- n Bittivirheen havaitsemiseen
- n UDP ei yritä toipua, hävittää  
virheellisen segmentin

## n Data

- n Pitkä sanoma pilkottuna useaksi segmentiksi

## n IP-osoitteet vasta verkkokerroksen otsakkeessa

- n Näitä tarvitaan reitityksessä



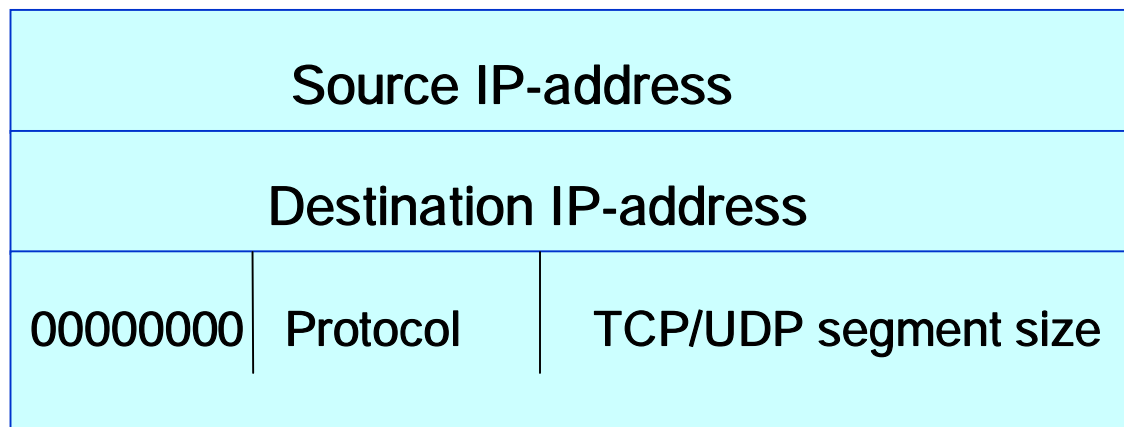
UDP-otsake





## Kuljetuskerroksen pseudo-otsake

- n Käyttö vain isäntäkoneen sisäisesti. Ei lähetetä verkkoon.
  - n UDP laskee tarkistussumman otsakkeelle, datalle ja ns. **pseudo-otsakkeelle**, joka sisältää IP-otsakkeen tietoja
  - n Varmistus, että segmentti on tullut oikeaan koneeseen ja oikeaan porttiin



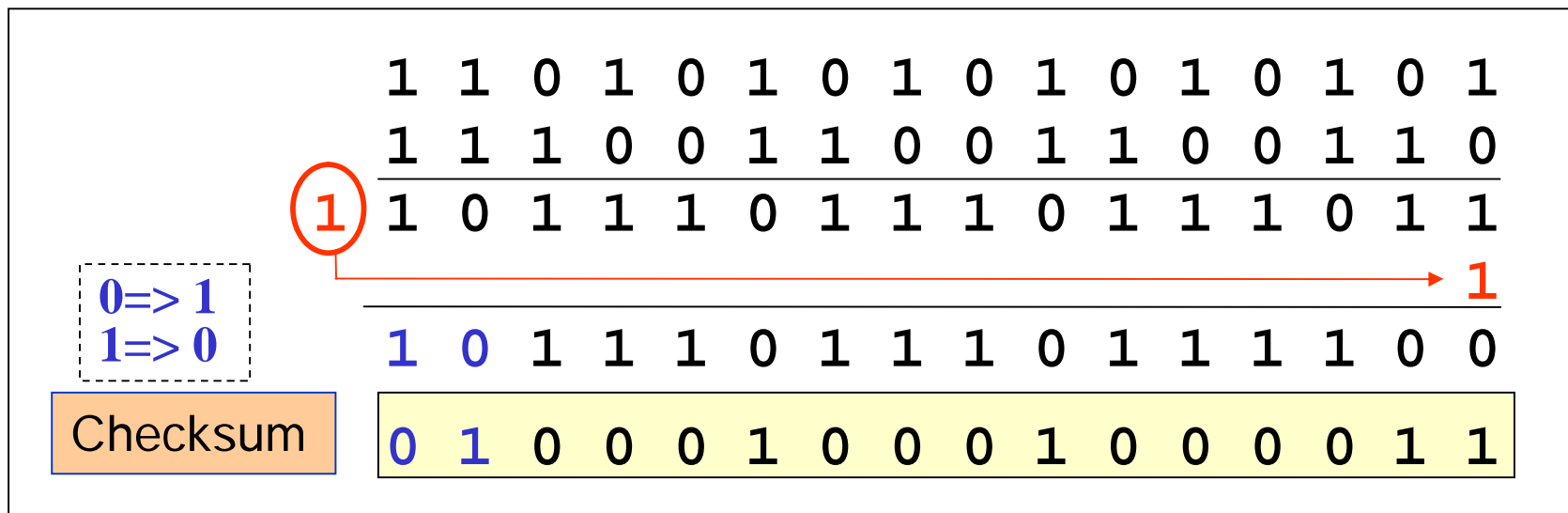
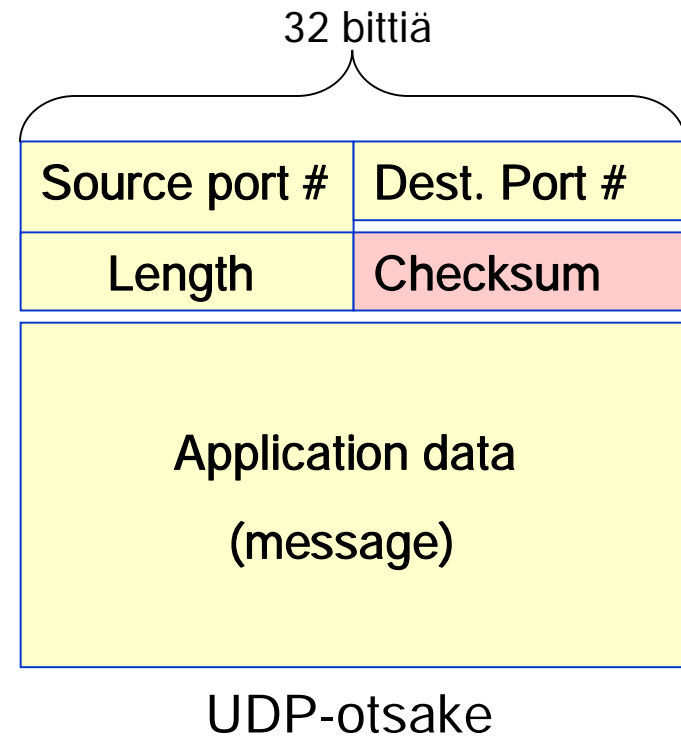
# UDP-tarkistussumma

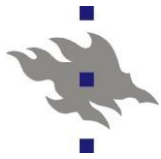
## n Lähetys

- n Summaa 16 bitin kokonaisuudet (otsake + pseudo-otsake mukana), ylivuotobitit lasketaan mukaan, talleta **yhden komplementtina**

## n Vastaanotto

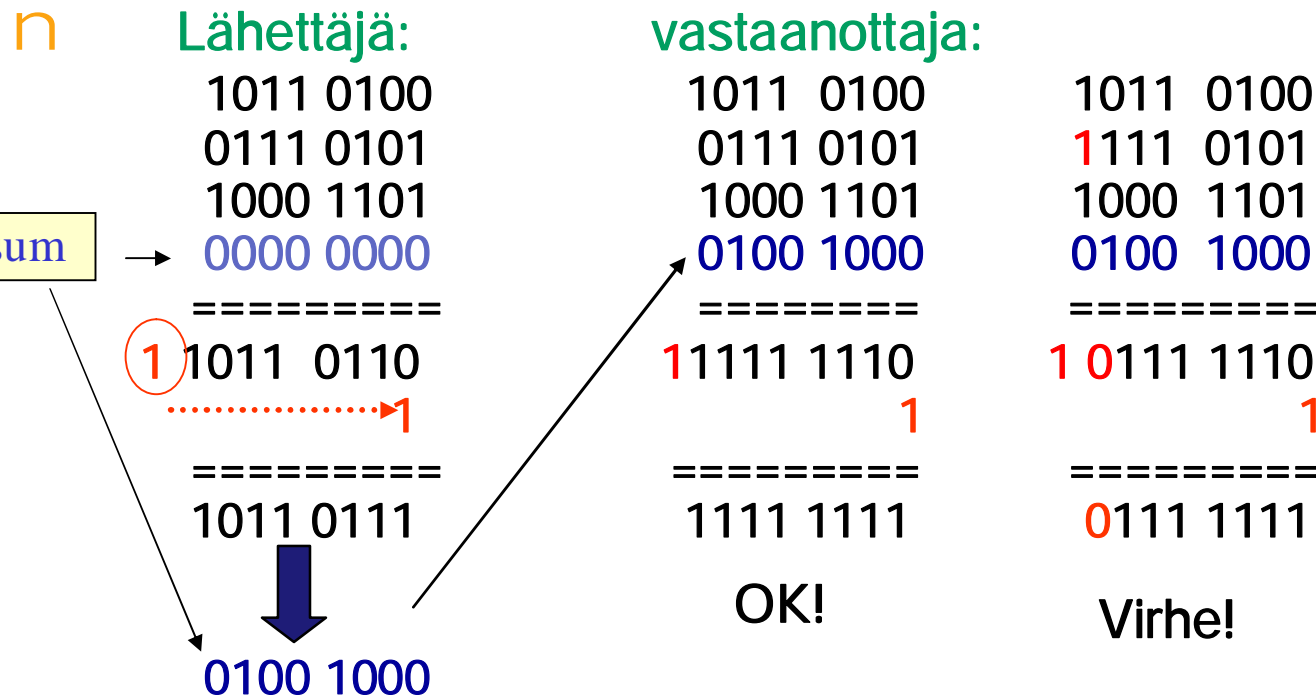
- n Summaa 16 b kokonaisuudet (myös tarkistussumma).
- n Jos tuloksena on 16 ykköstä, niin OK!





# Esimerkki

**n** Lasketaan tarkistussumma kolmen tavun mittaiselle sanomalle (tässä vain 8 bitin mittaisena):



OK!

Virhe!

Yhden komplementti



## Miksi UDP-tarkistussumma

- n Kaikki linkkikerrokset eivät suorita tarkistuksia!

- n Ethernet huolehtii kyllä

- n UDP-tarkistussumma ei ole kovin tehokas havaitsemaan virheitä!

- n UDP ei yritä toipua virheistä!

- n Jotkut toteutukset voivat tuhota virheellisen segmentin

- n Jotkut antavat sen sovellukselle varoituksen kera

- n Lisärasite?

- n Ei tarvitse käyttää, jos ei halua. Tällöin lähettäjä laittaa tarkistusummaksi pelkkiä nolliä



## Tehtäviä:

n Lähetetään 10 tavun viesti UDP:llä.

n Miten kauan kestää lähettäminen, jos lähetyksenopeus on 56 kbps?

$$10 \text{ tavua} + 8 \text{ tavua} = 18 * 8 \text{ b} = 144 \text{ bittiä}$$

$$144 \text{ b} / 56\,000 \text{ b/s} = 2.57 \text{ ms}$$

n Miten suuri on etenemisviive, jos etäisyys lähettäjältä vastaanottajalle on 1000 km?

$$1000 \text{ km} / 200\,000 \text{ km/s} = 5 \text{ ms}$$

n Miten suuri on UDP-otsakkeen aiheuttama yleisrasite (overhead)?

$$8/18 = 0.44 \text{ eli } 44 \%$$

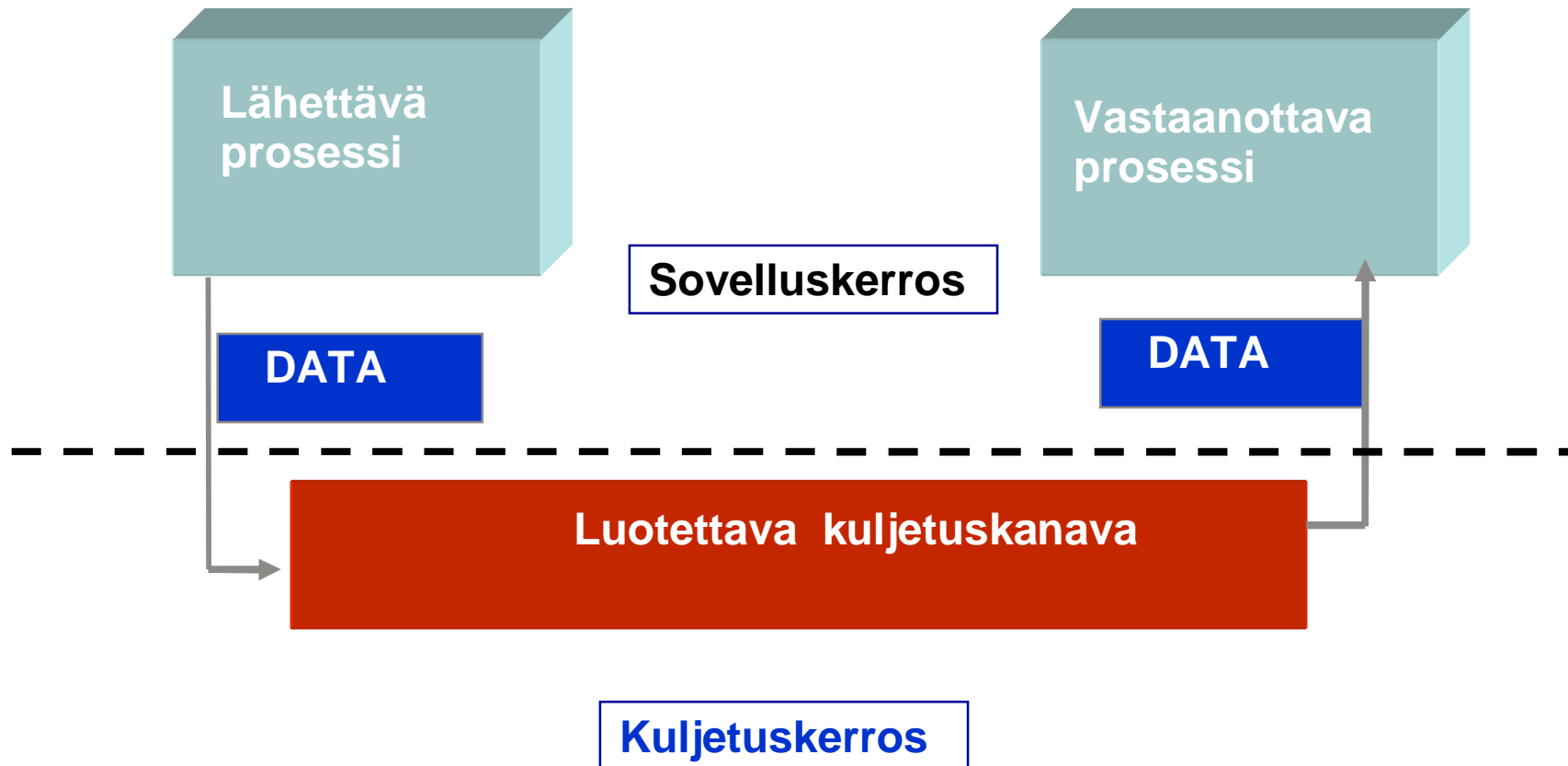


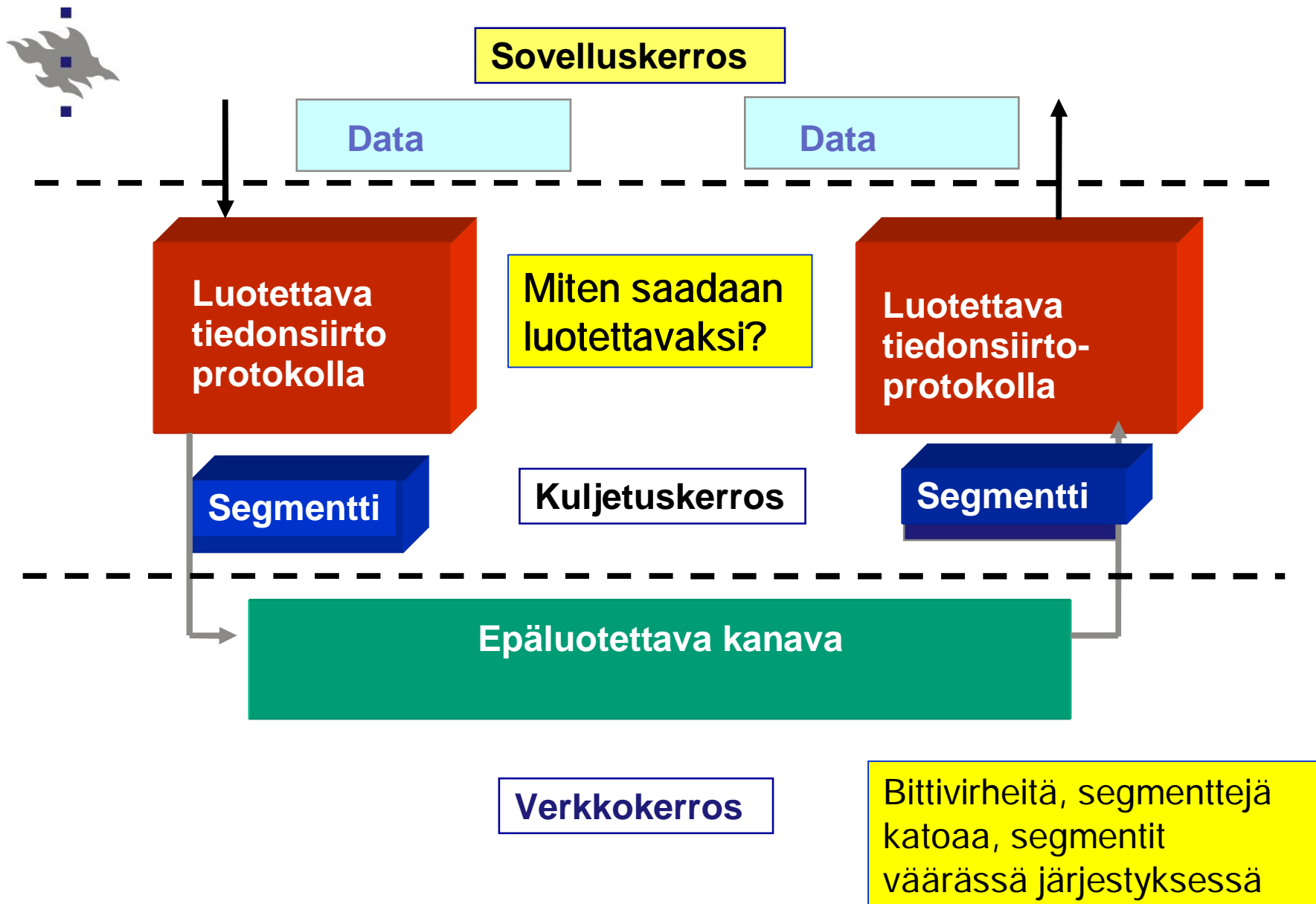
## Kuljetuskerros

# Luotettavan kuljetuspalvelun periaatteet

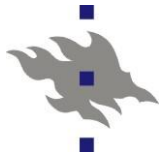


# Luotettava tiedonsiirto



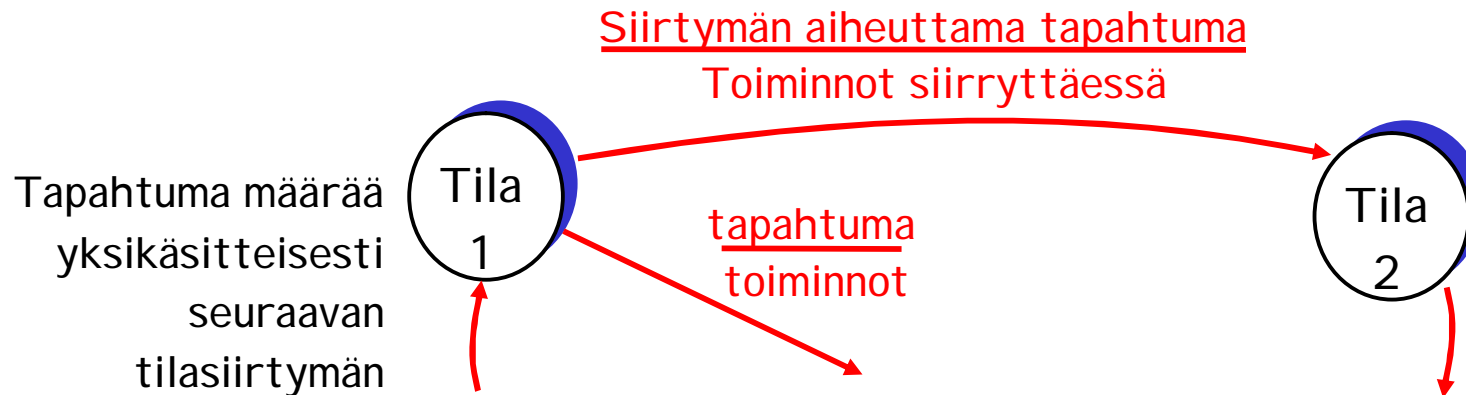






# Kuinka saada luotettavaksi?

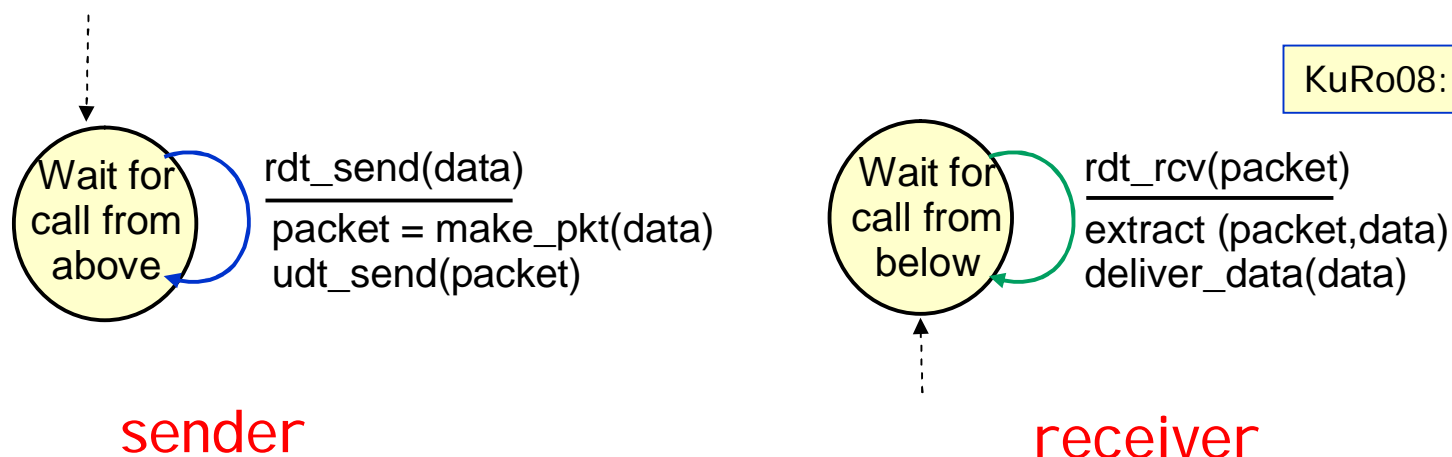
- r Tarkastellaan yleisesti luotettavan tiedonsiirron ongelmia ja erilaisia ratkaisuyrityksiä
  - r Edeten ideaalitalanteesta yhä ongelmaisempaan
  - r Käyttäen äärellisiä tila-automaatteja lähettäjän ja vastaanottajan toiminnan kuvaamiseksi

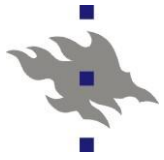




## Versio rdt1.0: Ideaalitilanne

- r Oletus: **siirtokanava on täysin luotettava**
  - r Ei bittivirheitä, ei katoavia paketteja, ei epäjärjestystä
- r Luotettava kuljetuspalvelu =
  - r Lähettäjä kirjoittaa datan kanavaan,
  - r Vastaanottaja lukee datan kanavasta





## Versio rdt2.0: Bittivirheitä

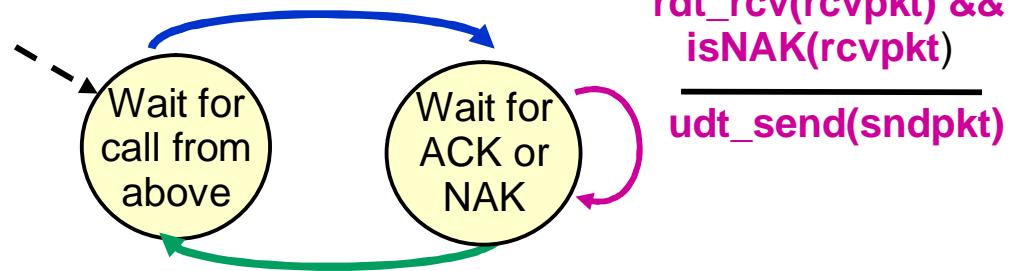
- r Oletus: Voi olla **vain bittivirheitä**
  - r Bitti voi kääntyä siirron aikana
  - r **Siirto ei kadota paketteja**
- r **Kuinka toipua?**
  - r Kuittaukset: vastaanottaja kuittaa **ACK**:lla virheettömän paketin,
  - r **NAK**-kuittaus, jos paketti on virheellinen + hylkää
  - r Jos NAK, niin lähettäjä lähettää paketin uudelleen
- r **Luotettava kuljetuspalvelu**
  - r Virheen huomaaminen: **tarkistussumma**
  - r Palaute vastaanottajalta: **kuittaussanoma** (ACK / NAK)
  - r **Uudelleenlähetys**: dataa puskuroitava
- r **Stop-and-wait-protokolla**
  - r Lähettäjä odottaa kuittausta ennenkuin lähettää seuraavan



## rdt2.0:

sender

rdt\_send(data)  
snpkt = make\_pkt(data, **checksum**)  
udt\_send(sndpkt)

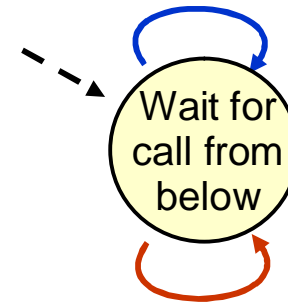


rdt\_rcv(rcvpkt) && isACK(rcvpkt)

Λ

receiver

rdt\_rcv(rcvpkt) && corrupt(rcvpkt)  
udt\_send(NAK)



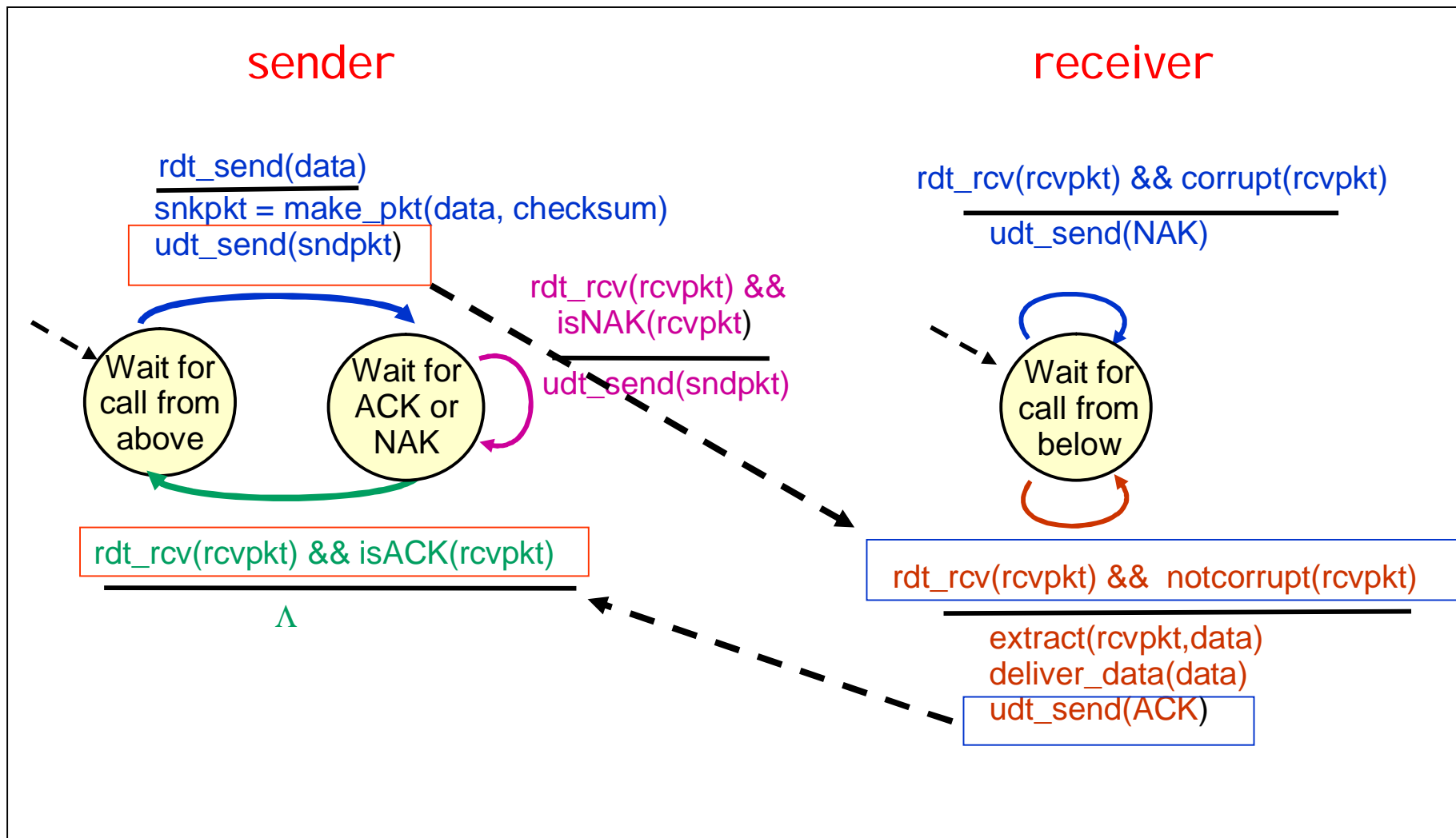
rdt\_rcv(rcvpkt) && notcorrupt(rcvpkt)

extract(rcvpkt,data)  
deliver\_data(data)  
udt\_send(ACK)

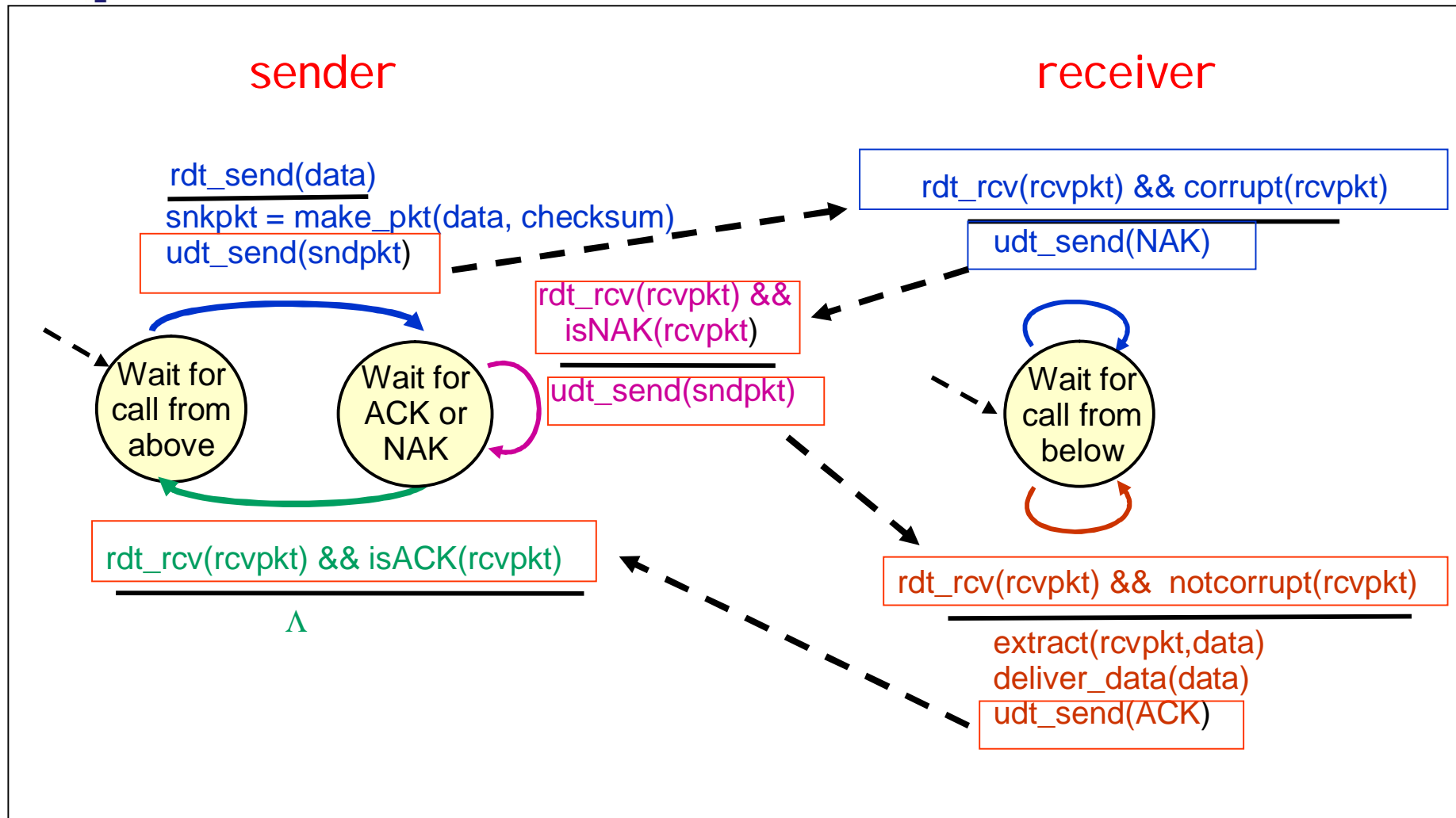
KuRo08: Fig. 3.10



## rdt2.0: Toiminta, kun ei ole virhettä



## rdt2.0: Toiminta virhetilaneessa





## rdt2.0: Missä voi mennä väärin?

- r Ei toimi, jos ACK /NAK -bitit korruptoituvat!
  - r Onko OK vai ei?
- r Korjaus:ACK/NAK-paketteihin tarkistusbitit, jotta virhe huomataan
- r Entä toipuminen?
  - r Jos jos kuittauksessa virhe, uudelleenlähetetään paketti
  - r Uudelleenlähetys voi tuottaa **kaksoiskappaleen, jotka on huomattava ja hylättävä**
- r => **Paketteihin järjestysnumero**
  - r Kaksoiskappale: jos sama numero
  - r Vastaanottaja kuittaa normaalisti, mutta ei anna sovellukselle



## Versio rdt2.1: enemmän bittivirheitä

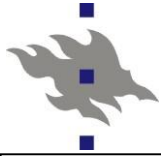
### Lähettäjä:

- n Lisää pakettiin järjestysnumeron  
(numerot 0 ja 1 riittävät tässä protokollassa. Miksi?)  
Ns. vuorottelevan bitin protokolla
- n Tarkista, että ACK/NAK ei ole korruptoitunut  
Tilakaaviossa nyt kaksinkertaisesti tiloja  
Kaavion tilan 'muistettava' paketin numero
- n Säilytä kopio lähetetystä paketista

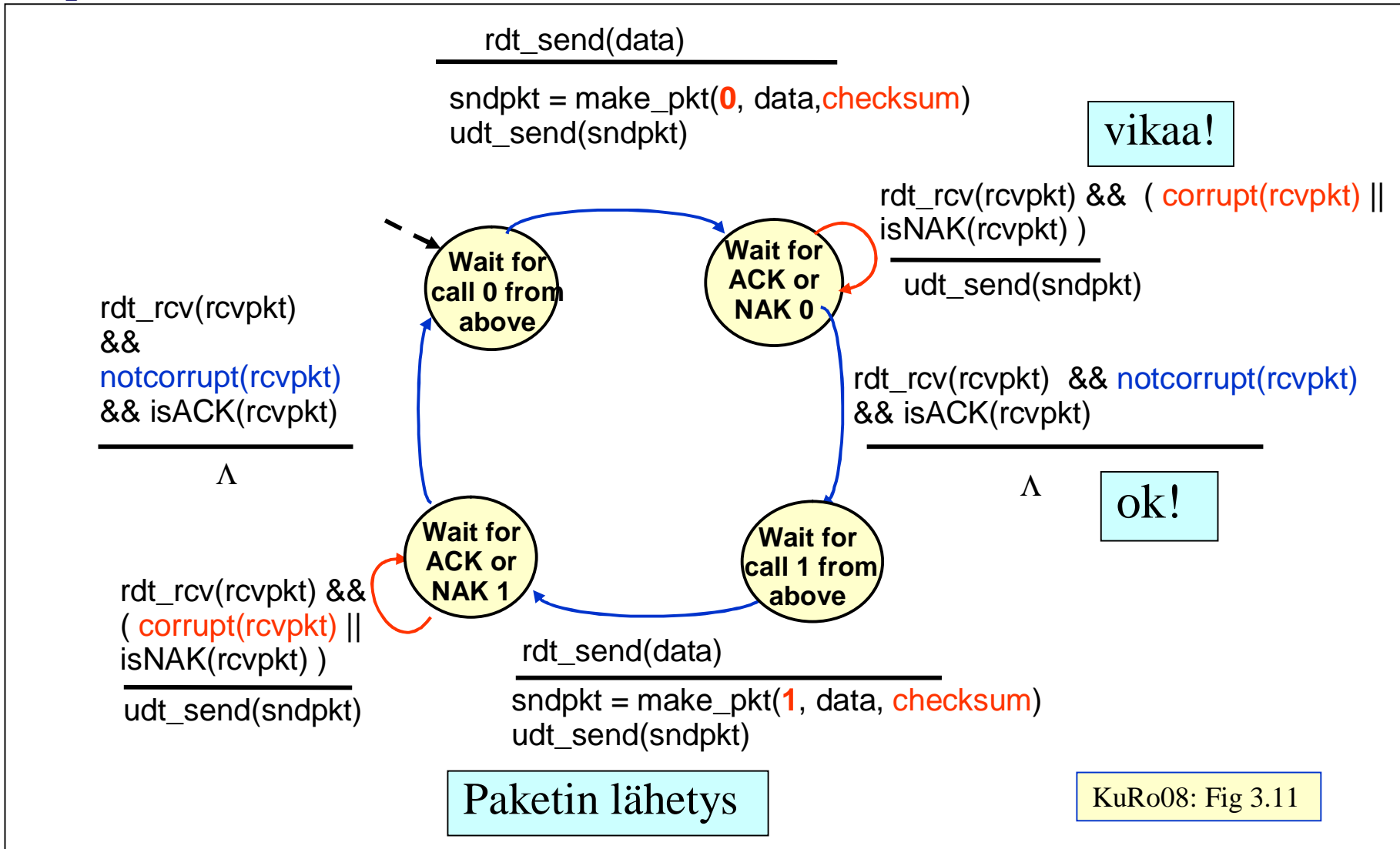
### Vastaanottaja:

- n Tarkista, ettei ole kaksoiskappale
  - n Kaavion tilan 'muistettava' seuraavan paketin numero: 0 vai 1
- n Myös duplikaatti kuitattava, koska ei tiedä menikö edellinen kuittaus perille

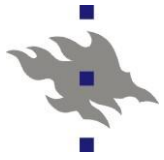




# rdt2.1: Lähettäjän tilakaavio



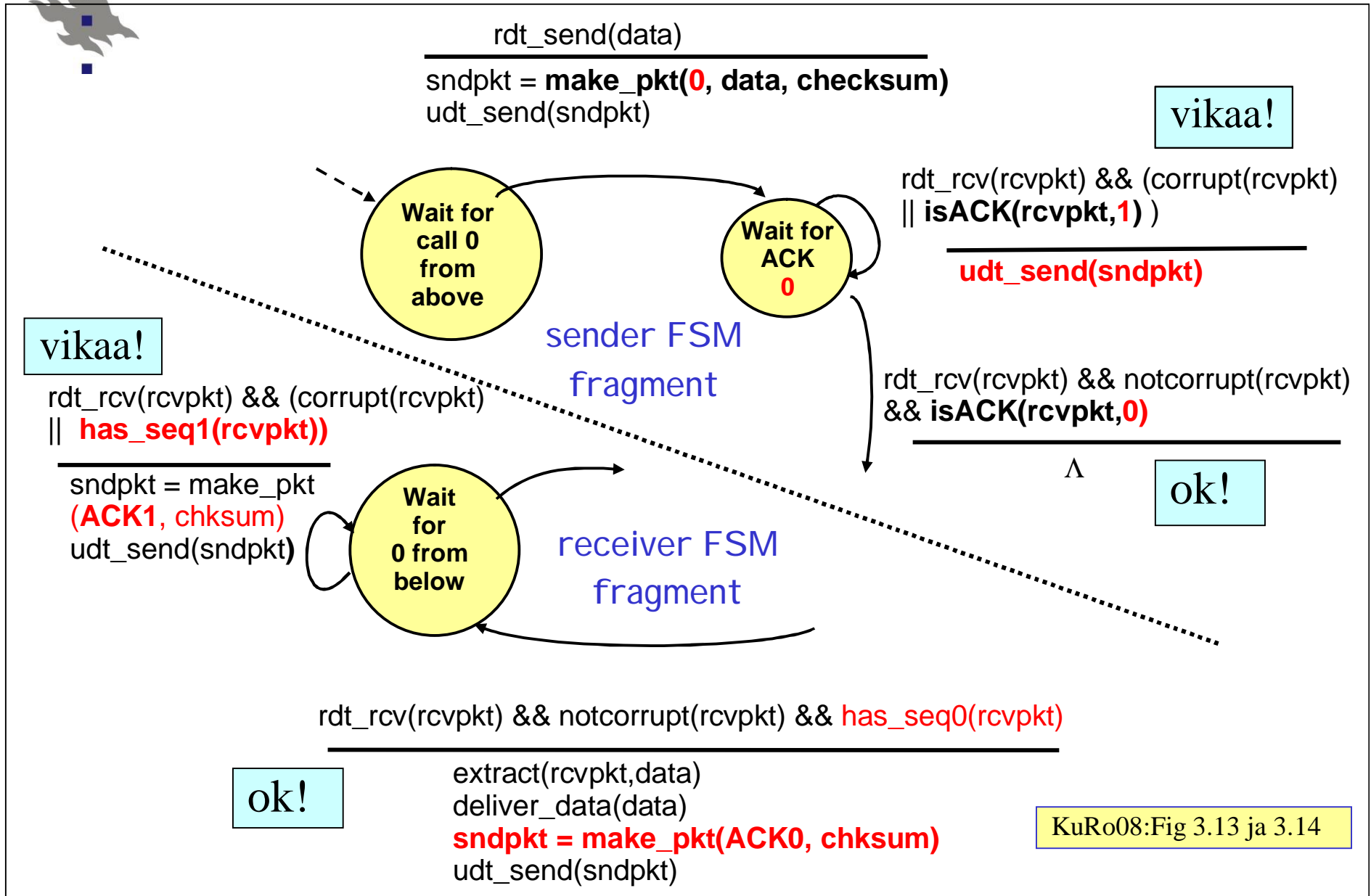




## Versio rdt2.2: vain ACK-kuittaus käytössä

- r Sama toiminnallisuus kuin edellä
- r Käyttää vain ACK-kuittauksia
  - r Vastaanottaja kuittaa viimeksi kunnossa saamansa paketin
  - r Kuittaukseen on liitettävä kuitattavan paketin numero
- r Jos samalle paketille (nro X) tulee useita ACK-kuittauksia (*duplicate ACK*), niin sitä seuraava paketti (nro X+1) joko puuttuu tai on virheellinen
  - r ~NAK-kuittaus
  - r Lähetetään uudelleen sanoma X+1

# rdt2.2

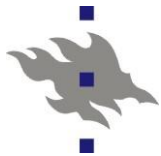


KuRo08:Fig 3.13 ja 3.14

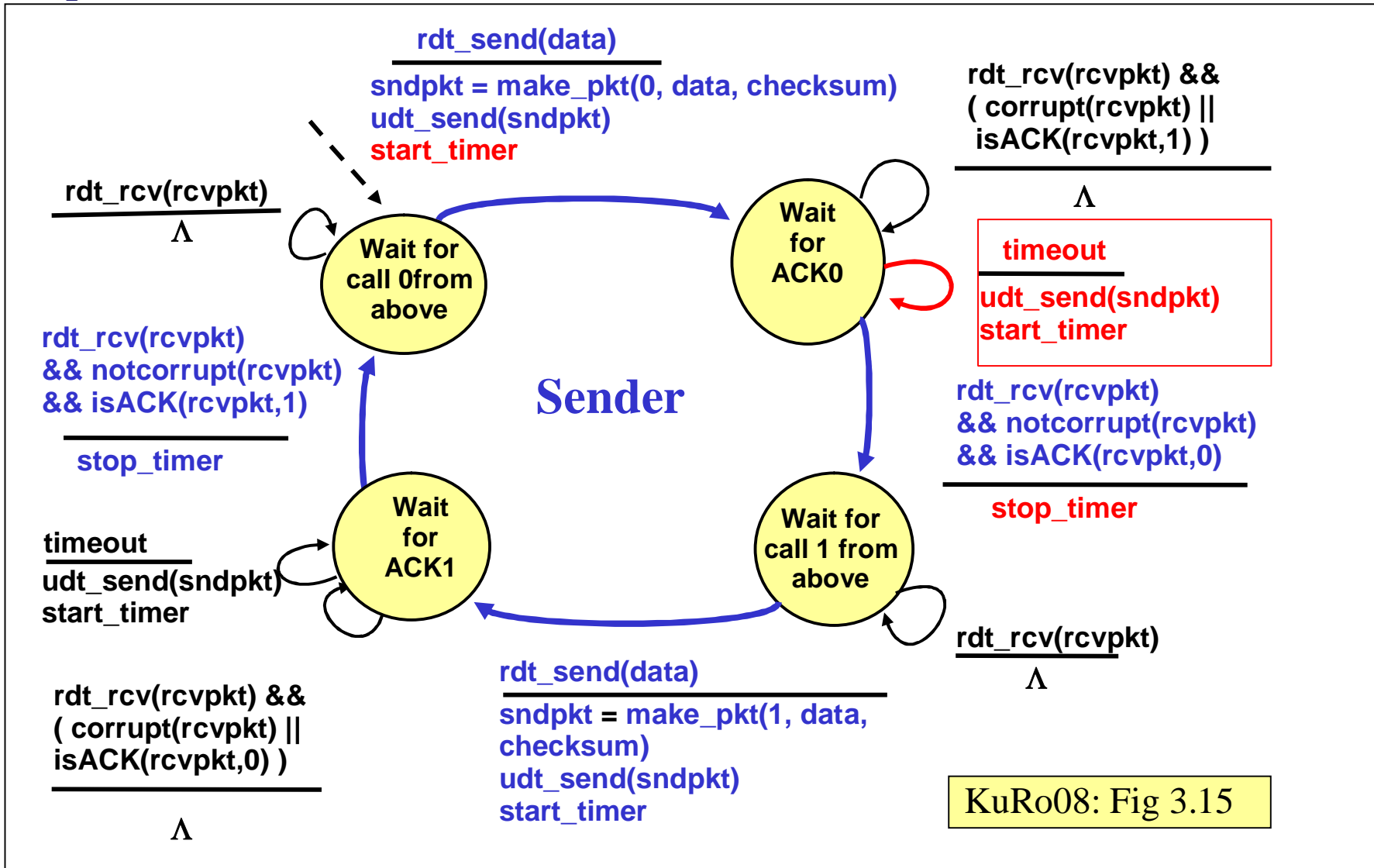


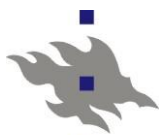
## Versio rdt3.0: paketteja voi kadota!!

- Q **Oletus: Siirtokanava voi kadottaa paketteja**  
Sekä datapaketteja että kuittauspaketteja voi kadota.  
Tarkistussumma, pakettinumero, ACK eivät vielä riitä! **Miksi?**
- Q Lähettäjä odottaa jonkin aikaa kuittausta  
Jos ei saavu, lähettää paketin uudelleen  
**Ajastin** laukaisee uudelleenlähetyksen
- Q Jos paketti (data / kuittaus) kuitenkin vain viivästyy eikä olekaan kadonnut  
Syntyy duplikaatti, joka havaitaan sanomanumeroinnin avulla  
Kuittauksessa mukana kuitatun paketin numero

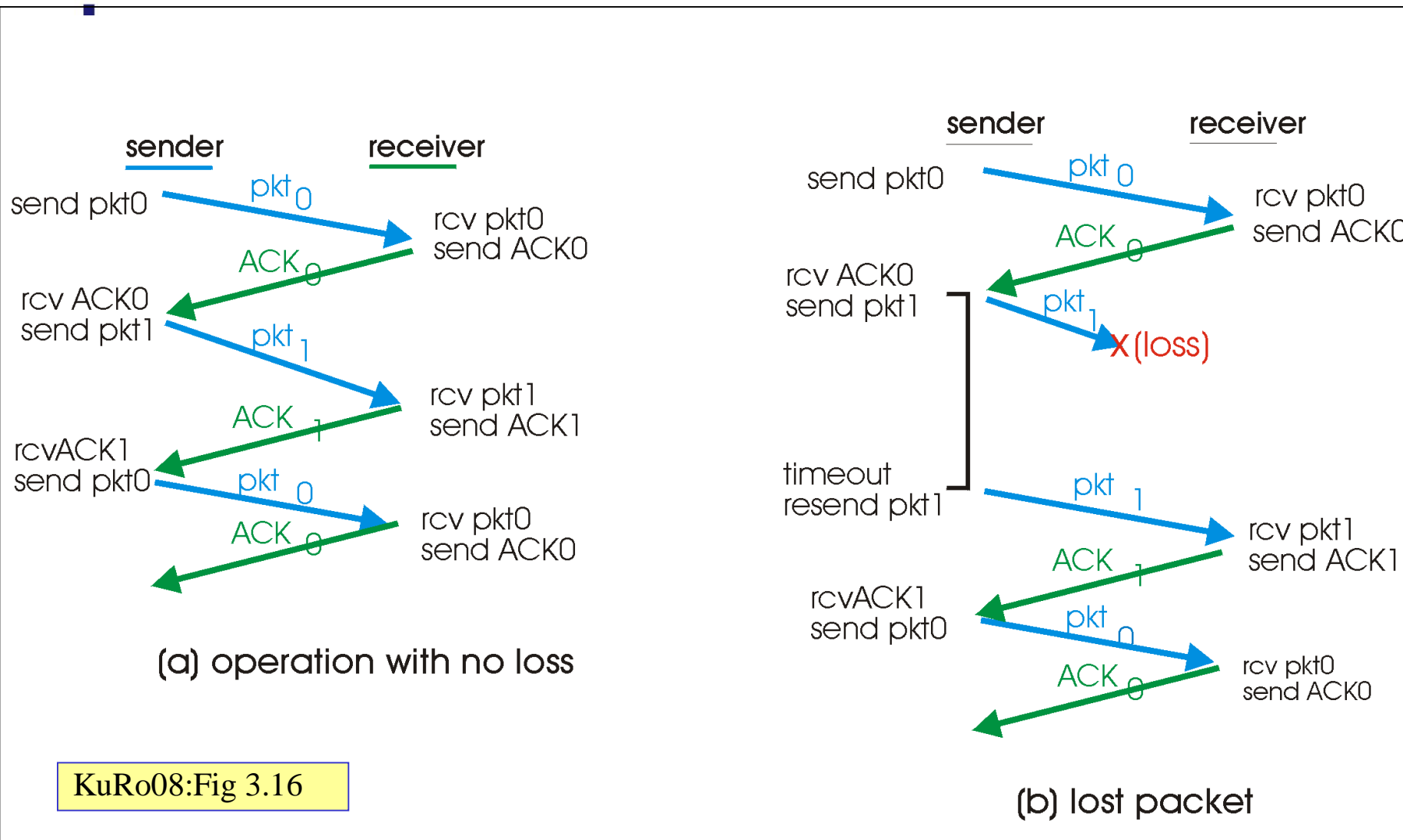


# rdt3.0



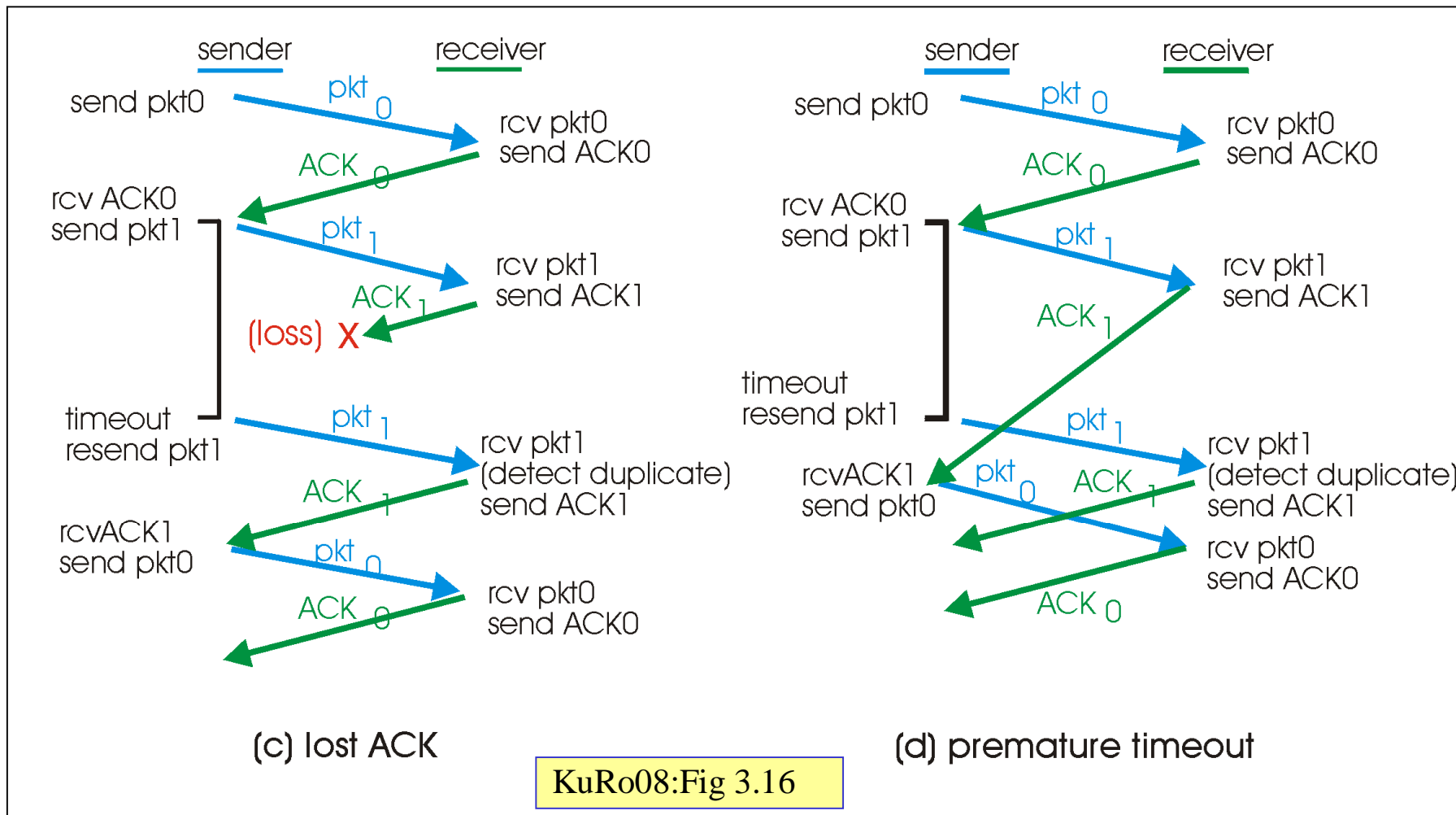


# rdt3.0 toiminnassa





# rdt3.0 toiminnassa







## rdt3.0: Tehokkuus?

Esim: 1 Gbps linkki, 15 ms päästä-päähän etenemisviive eli  
RTT = 30 ms, 1 KB:n paketti

$$T_{\text{transmit}} = \frac{L \text{ (packet length in bits)}}{R \text{ (transmission rate, bps)}} = \frac{8\text{kb/pkt}}{10^{**}9 \text{ b/sec}} = 8 \text{ microsec}$$
$$U_{\text{sender}} = \frac{L / R}{RTT + L / R} = \frac{.008}{30.008} = 0.00027$$

**Käyttöaste** (utilization): se osa kokonaisajasta, jolloin lähettäjä lähettää

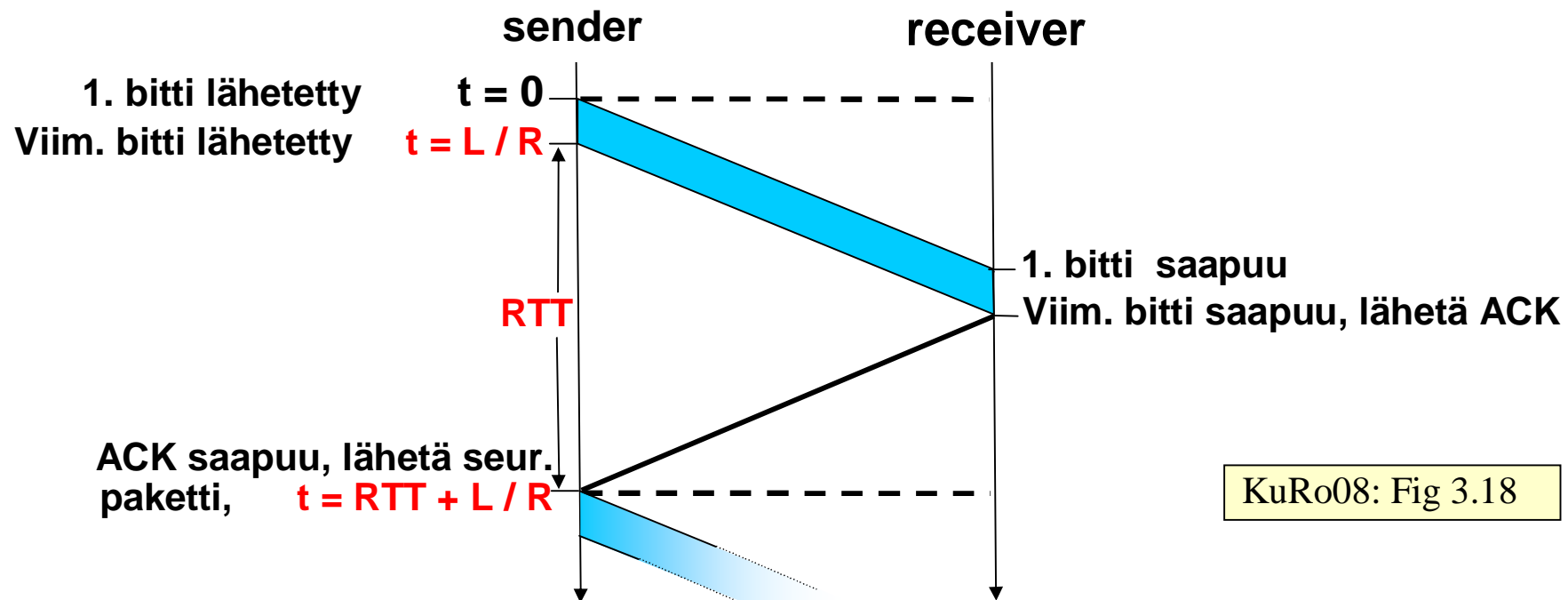
m 1KB paketti 30 ms:n välein -> 33kB/s nopeus 1 Gbps linkillä.

m Stop-and-wait-protokolla rajoittaa,

ei linkin kyky siirtää dataa (linkin siirtonopeus)



## rdt3.0: stop-and-wait tehokkuus



KuRo08: Fig 3.18

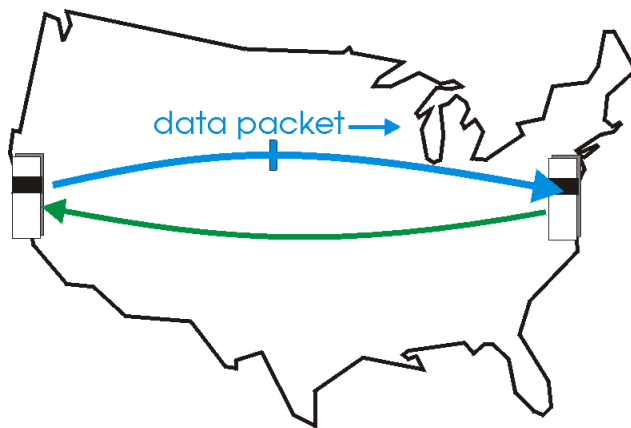
$$U_{\text{sender}} = \frac{L / R}{RTT + L / R} = \frac{.008}{30.008} = 0.00027$$



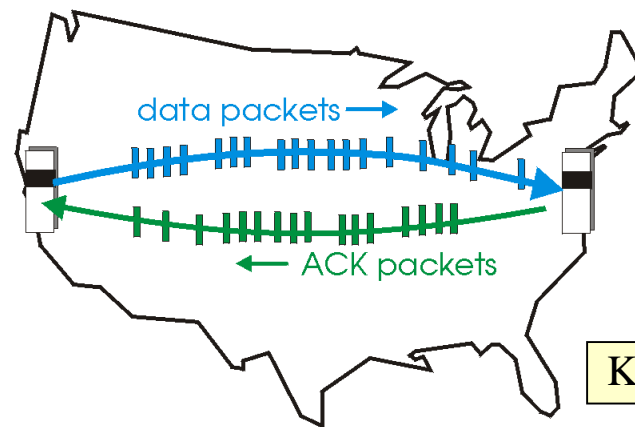
## rdtX.X: Liukuhihnaprotokollat

Lähettäjä saa lähettää useita paketteja, vaikka ei ole saanut kuittauksia edeltäviin

- n Numerointi (0,1) ei enää riitä, lisää numeroita tarvitaan
- n Tarvitaan puskurointia molemmissa päissä

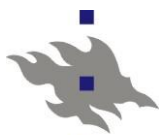


(a) a stop-and-wait protocol in operation

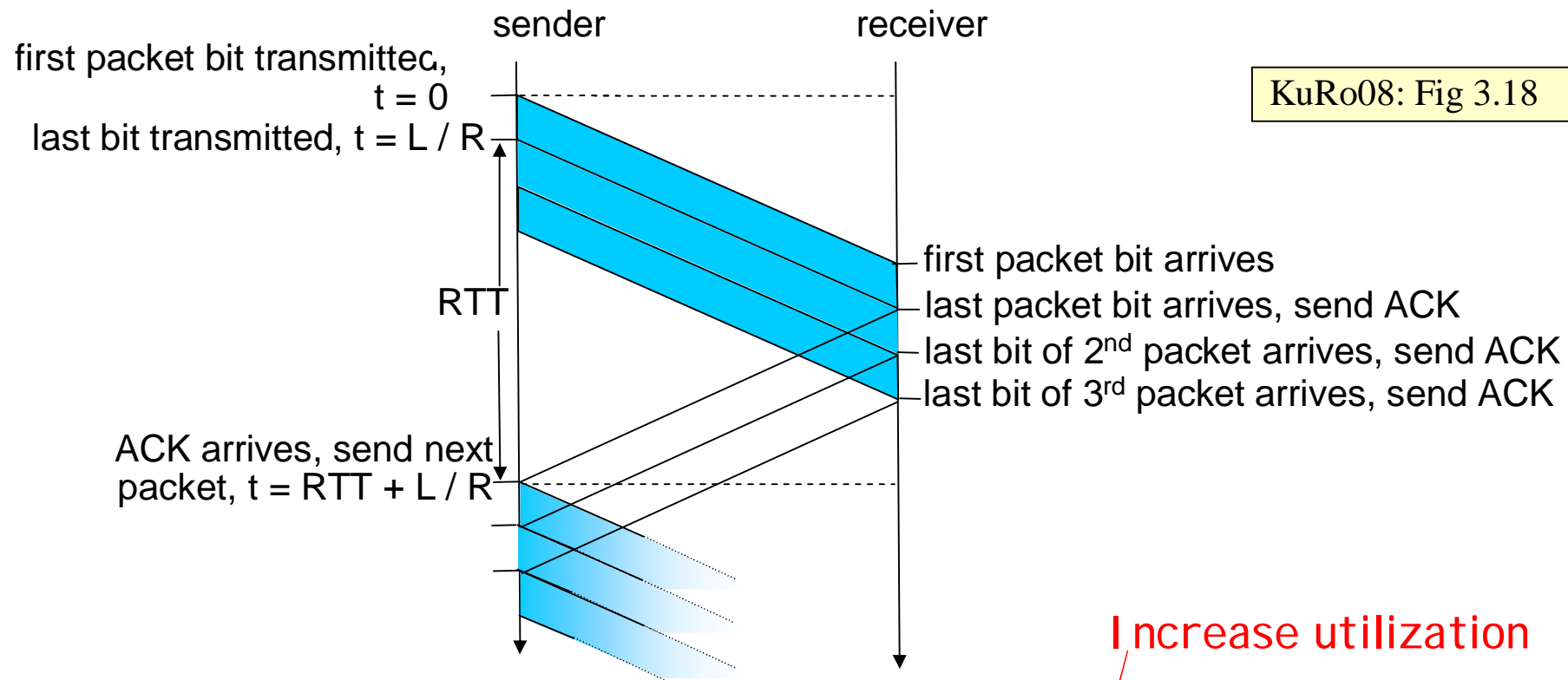


(b) a pipelined protocol in operation

KuRo08: Fig 3.17



# Liukuhhnoitus: käyttöasteen kasvattaminen

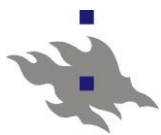


KuRo08: Fig 3.18

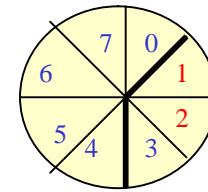
Increase utilization  
by a factor of 3!

$$U_{\text{sender}} = \frac{3 * L / R}{RTT + L / R} = \frac{.024}{30.008} = 0.0008$$

(0.00028)



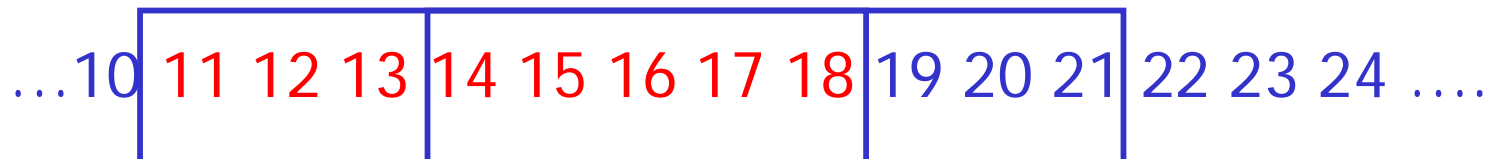
## Liukuva ikkuna (sliding window)



### n Säätää pakettien lähettämistä ja vastaanottoa

n Kertoo millä pakettinumeroilla on lähetetty/vastaanotettu, mistä saatu/lähetetty kuittaukset ja millä numeroilla voi vielä lähettää/vastaanottaa paketteja

n Ikkunan koko riippuu yhteyden tyypistä ja puskurien koosta



### n Lähetysikkuna (sender window)

n Ikkunan koko = montako pakettia saa olla kuittaamatta

n Mitkä pakettinumerot on käytetty, mutta kuittaamatta

n Mitä pakettinumeroita voi vielä käyttää

n Lähettäjän on odotettava, jos kaikki ikkunan numerot on käytetty

n Kun kuittaus saapuu, ikkuna liukuu

n Seuraavat numerot tulevat luvallisiksi



## Liukuva ikkuna (2)

### n **Vastaanottoikkuna** (receiver window)

- n Mitkä pakettinumerot otettu vastaan, mutta kuittaamatta
- n Mitä pakettinumeroita lähettäjä saa vielä käyttää eli mitkä hyväksytään

### n Jos saadussa paketissa on ikkunan viimeinen numero

- n Ikkuna pysäyttää pakettien lähetyksen vastapäätä
- n Ikkuna estää uusien pakettien vastaanoton

### n Paketin kuittaus liu'uttaa myös vastaanottajan ikkunaa

- n Hyväksytään uusia pakettinumeroita



## Kun ikkunan koko on 1

- n Vain yksi paketti kuittaamattomana

  - n One Bit Sliding Window –protokolla

  - n = stop-and-wait –protokolla

- n Pakettinumerot 0 ja 1 riittävät

- n ACK ilmoittaa

  - n Joko seuraavaksi odotetun paketin numeron (esim. TCP)

  - n Tai viimeksi vastaanotetun virheettömän paketin numeron

- n ACK sisältää paketin numeron

  - n Kuittausduplikaatti ei voi kuitata väärää paketteja



# Virhetilanteen käsittely

n Entä, kun huomataan virhe?

n Monta muuta pakettia jo matkalla!

n Pakettiin ei tule kuittausta

n Paketti katosi tai virheellinen

n Kuittaus katosi tai virheellinen

n => Ajastin laukeaa aikanaan

n "Go-Back-N" (paluu N:ään)

n Paketit uudelleenlähetetään virheellisestä lähtien

n Selective Repeat (Valikoiva toisto)

n Lähetetään vain virheettiset paketit





## Go-Back-N

n Vastaanottaja **hyväksyy paketit vain järjestyksessä**

- n Kuittaa järjestyksessä tulleen virheettömän paketin
- n Hylkää kaikki puuttuvan tai virheellisen paketin jälkeiset paketit eikä lähetä niistä kuittauksia

n Kun lähettäjä ei saa pakettiin kuittausta

- n Lähetysikkuna täyttyy ja estää uusien pakettien lähettämisen
- n Lähettäjän ajastimet laukeavat
- n Lähettäjä lähettää uudestaan kaikki viimeisen kuittauksen jälkeiset paketit
- n Näiden kuittaukset siirtävät taas lähetysikkunaa

**n Tehoton, jos paljon virheitä ja iso ikkuna**



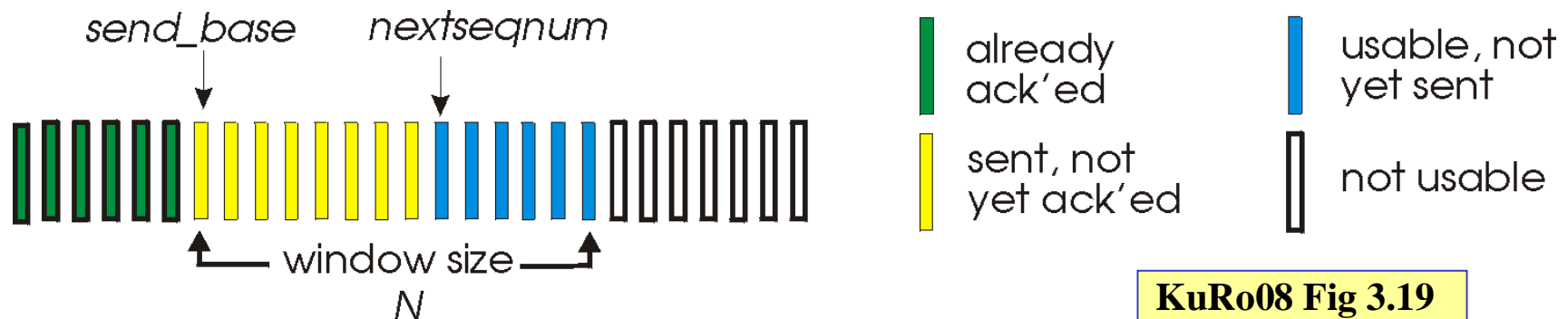
# Go-Back-N

## n Kumulatiivinen ACK

- n Lähetä ACK, jossa korkein järjestyksessä saadun kelvollisen paketin numero
- n Tämä kuittaa kaikki pienemmällä numerolla lähetetyt paketit

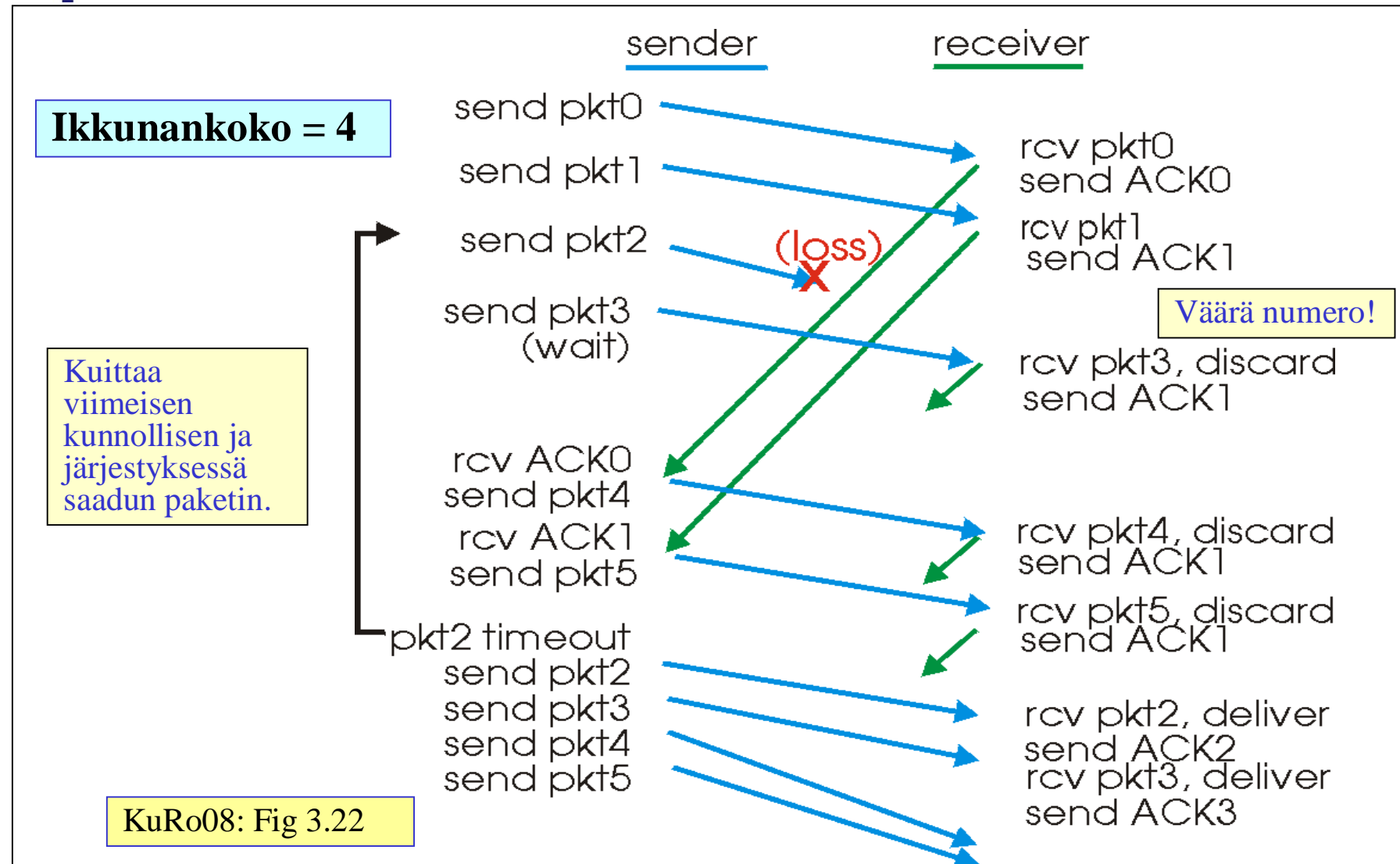
## n Jos välistä puuttuu paketti

- n Lähetä uudestaan ACK, jossa korkein järjestyksessä saadun paketin numero (~ NAK)
- n **Tuplakuittaus** (duplicate ACK)
- n Parannus: => nopeampi reagointi puuttuvaan





# Go-Back-N: Esimerkki



# Go-Back-N: lähettäjän tilakaavio

Lähetysoyntö sovellukselta: lähetetään, jos ikkuna sallii

rdt\_send(data)

```

if (nextseqnum < base+N) {
    sndpkt[nextseqnum] = make_pkt(nextseqnum,data,chksum)
    udt_send(sndpkt[nextseqnum])
    if (base == nextseqnum)
        start_timer
    nextseqnum++
}
else refuse_data(data)
    
```

Tässä vain yksi ajastin!

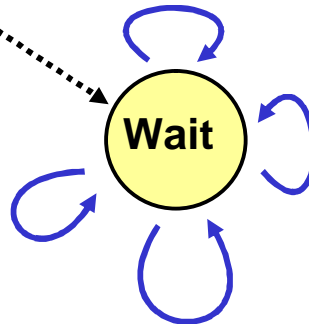
alkuarvot

$\Lambda$   
base=1  
nextseqnum=1

Ajastin laukeaa, lähetä kaikki kuittaamattomat uudestaan

rdt\_rcv(rcvpkt) && corrupt(rcvpkt)

$\Lambda$



timeout

```

start_timer
udt_send(sndpkt[base])
udt_send(sndpkt[base+1])
...
udt_send(sndpkt[nextseqnum-1])
    
```

Korruptoitunut kuittaus hylätään

```

rdt_rcv(rcvpkt) &&
notcorrupt(rcvpkt)
base = getacknum(rcvpkt)+1
If (base == nextseqnum)
    stop_timer
else start_timer
    
```

Kuittaus siirtää ikkunaa

KuRo08: Fig 3.20

Sender



## Go-Back-N: Vastaanottajan tilakaavio

receiver

```
rdt_rcv(rcvpkt) && notcorrupt(rcvpkt)
&&
hasseqnum(rcvpkt,expectedseqnum)
```

Paketti ok ja  
odotettu numero

```
extract(rcvpkt,data)
deliver_data(data)
sndpkt=make_pkt(expectedseqnum,ACK,chksum)
udt_send(sndpkt)
expectedseqnum++
```

Data sovellukselle,  
kuittaus lähettäjälle  
Seuraava pnumero++

Λ

Wait

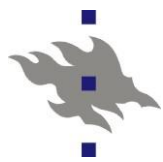
default  
udt\_send(sndpkt)

Muuten sama  
kuittaus  
uudestaan

```
expectedseqnum=1  
sndpkt =  
make_pkt(expectedseqnum,ACK,chksum)
```

alkuarvot

KuRo08: Fig 3.21



## Valikoiva toisto (Selective Repeat)

### n Valikoiva uudelleenlähetys

- n Lähetä uudelleen vain virheellinen /puuttuva paketti

### n Kuittaus jokaiselle kelvolliselle paketille

### n Paketit sovellukselle oikeassa järjestyksessä

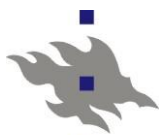
- n Vastaanottajalla oltava puskuritilaa pakettien järjestämiseen

### n Jos lähettäjä ei saa kuittausta paketista

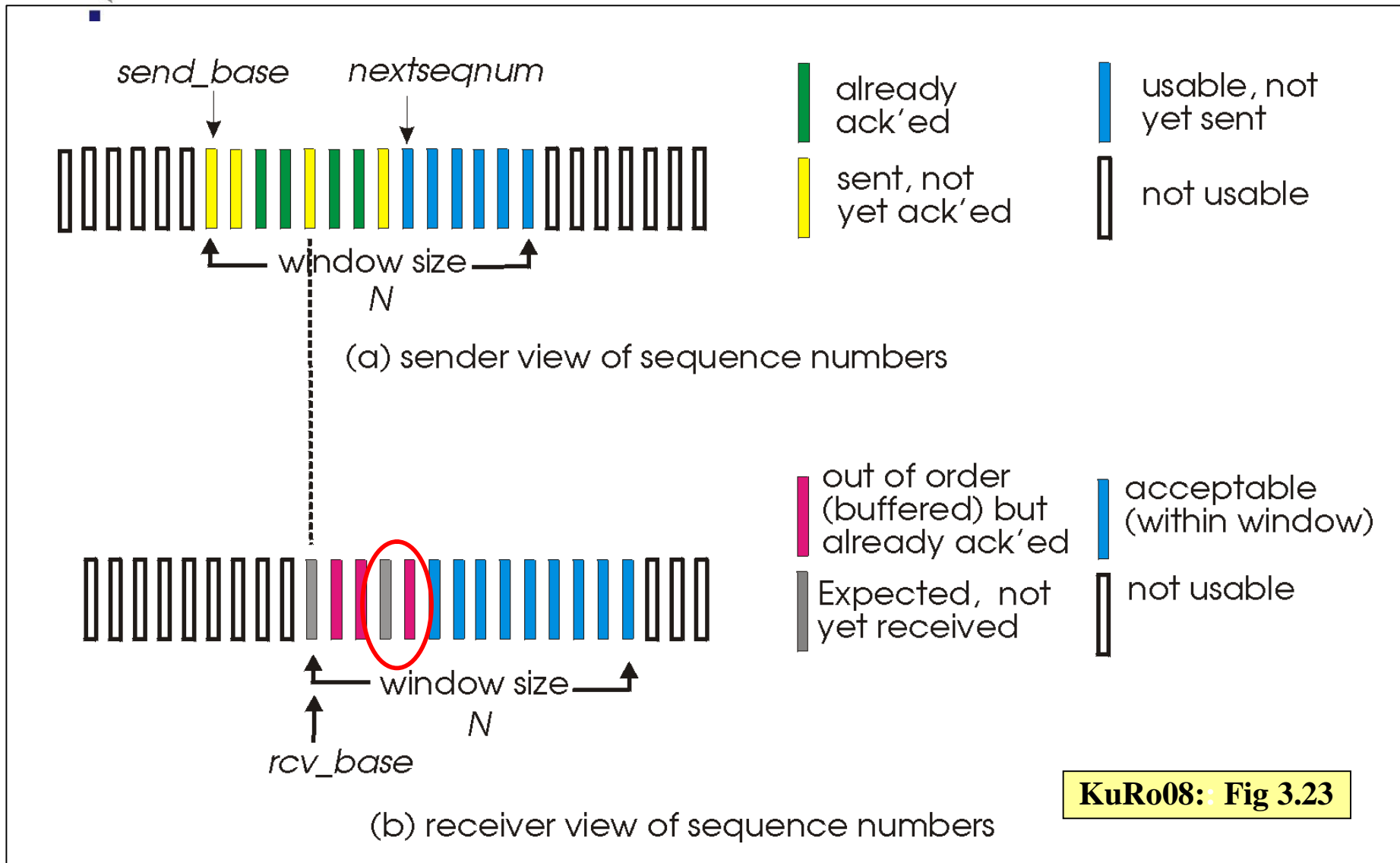
- n Lähetysikkunan täytyminen pysäyttää lähettämisen
- n Aikanaan ajastin laukeaa ja aiheuttaa uudelleenlähetyksen
- n Jokaisella paketilla on oma ajastin

### n Ikkuna liukuu nytkin tasaisesti

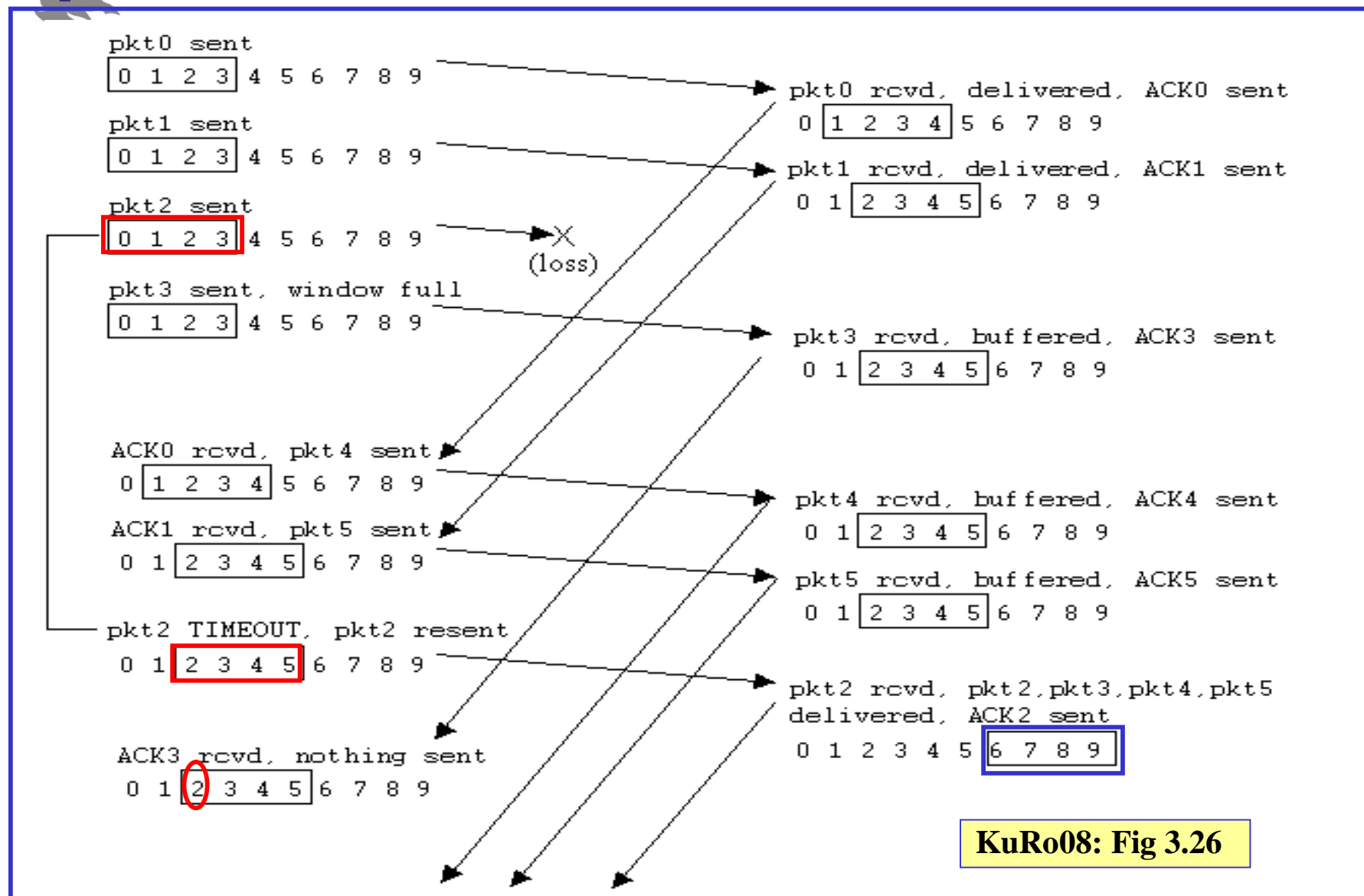
- n Yksi puuttuva kuittaus voi pysäyttää lähetyksen
- n Kun puuttuva paketti saatu, ikkuna liukuu kaikkien kuitattujen yli



# Valikoiva toisto



# Esimerkki



**KuRo08: Fig 3.26**







# Yhteenveto menetelmistä

n Ks. KuRo08 Table 3.1

n Tarkistussumma

n Ajastin

n Järjestysnumero

n Kuittaukset

n Positiiviset ACK, tuplakuittaukset

n Negatiiviset NAK

n Ikkunat, liukuhihnoitus

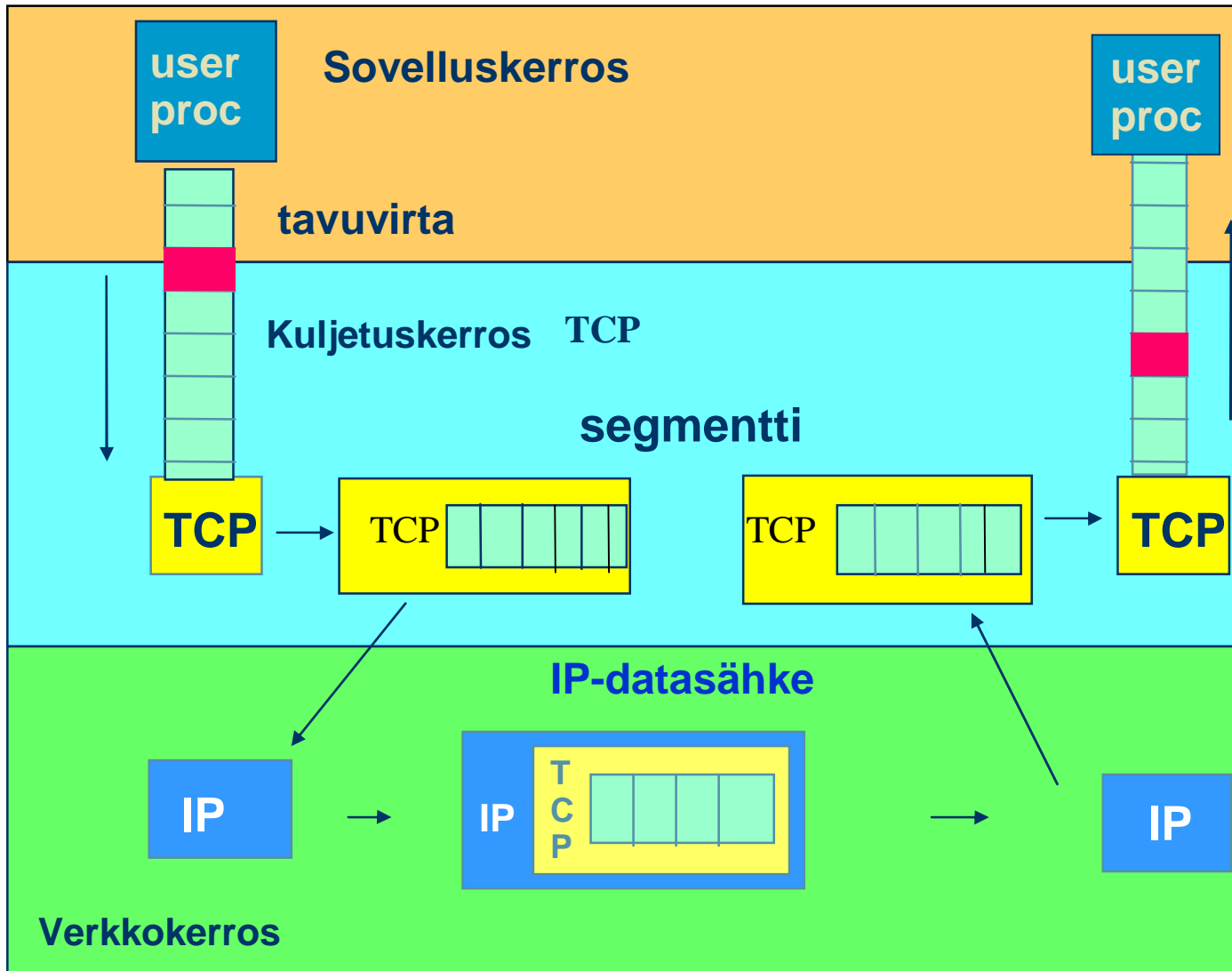
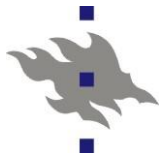


## Kuljetuskerros

# Yhteydellinen kuljetuspalvelu TCP

RFC 793, RFC 1122,  
RFC 1323, RFC 2018,  
RFC 2581

# TCP: prosessilta prosessille -tavuvirta





# TCP-protokolla

## n Päästä-päähän kuljetuspalvelu

- n Yksi lähettäjä, yksi vastaanottaja
- n Reitittimet eivät ole kiinnostuneita kuljetustason protokollasta

## n Yhteydellinen (connection-oriented)

- n Yhteydenmuodostus
- n Isäntäkoneissa: puskuritila, ikkunakoko, tavunumerointi
- n Yhteyden purku

## n Kaksisuuntainen (full duplex)

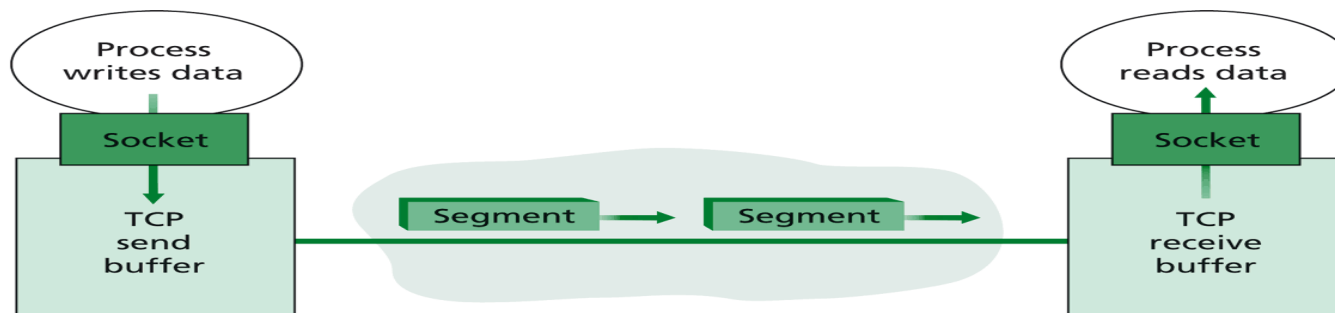
- n Yksi yhteys, jossa yhtä aikaa liikennettä molempiin suuntiin

## n Luotettava, järjestyksen säilyttävä tavuvirta

- n Ei sanomarajoja
- n Tavunumerointi
- n Kumulatiiviset kuittaukset

# TCP-protokolla

- n Vuonvalvonta, ruuhkanhallinta (-valvonta)
  - n Lähettäjä ei voi tukahduttaa vastaanottajaa eikä reitittämiä
- n Liukuvan ikkunan protokolla
  - n Vuonvalvonta ja ruuhkanhallinta vaikuttavat lähetyssikkunankokoon
- n Puskurointia molemmissa päissä
  - n Uudelleenlähetystä varten
  - n Jotta saadaan annettua sovellukselle järjestyksessä



**Figure 3.28** ♦ TCP send and receive buffers

# TCP-protokolla

## n Segmentillä maksimikoko

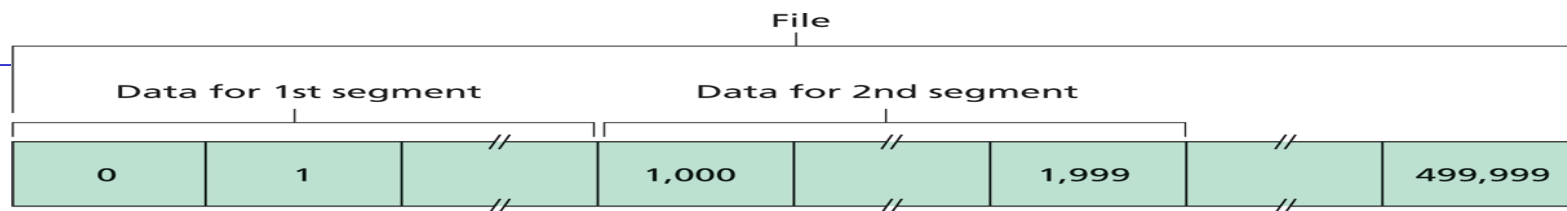
n MSS (maximum segment size) = paljonko dataa segmentissä

n Varmistaa, että tässä koneessa ei tarvita lisäpilkkomista paketeiksi

n Linkkikerroksen fyysiset ominaisuudet vaikuttavat MSS:n arvoon

n MTU (maximum transfer unit)

n Ethernet MTU = 1500 => MSS = 1460, sillä TCP:n ja IP:n osoitteet (20 + 20 tavua) vievät oman tilansa



**Figure 3.30** ♦ Dividing file data into TCP segments

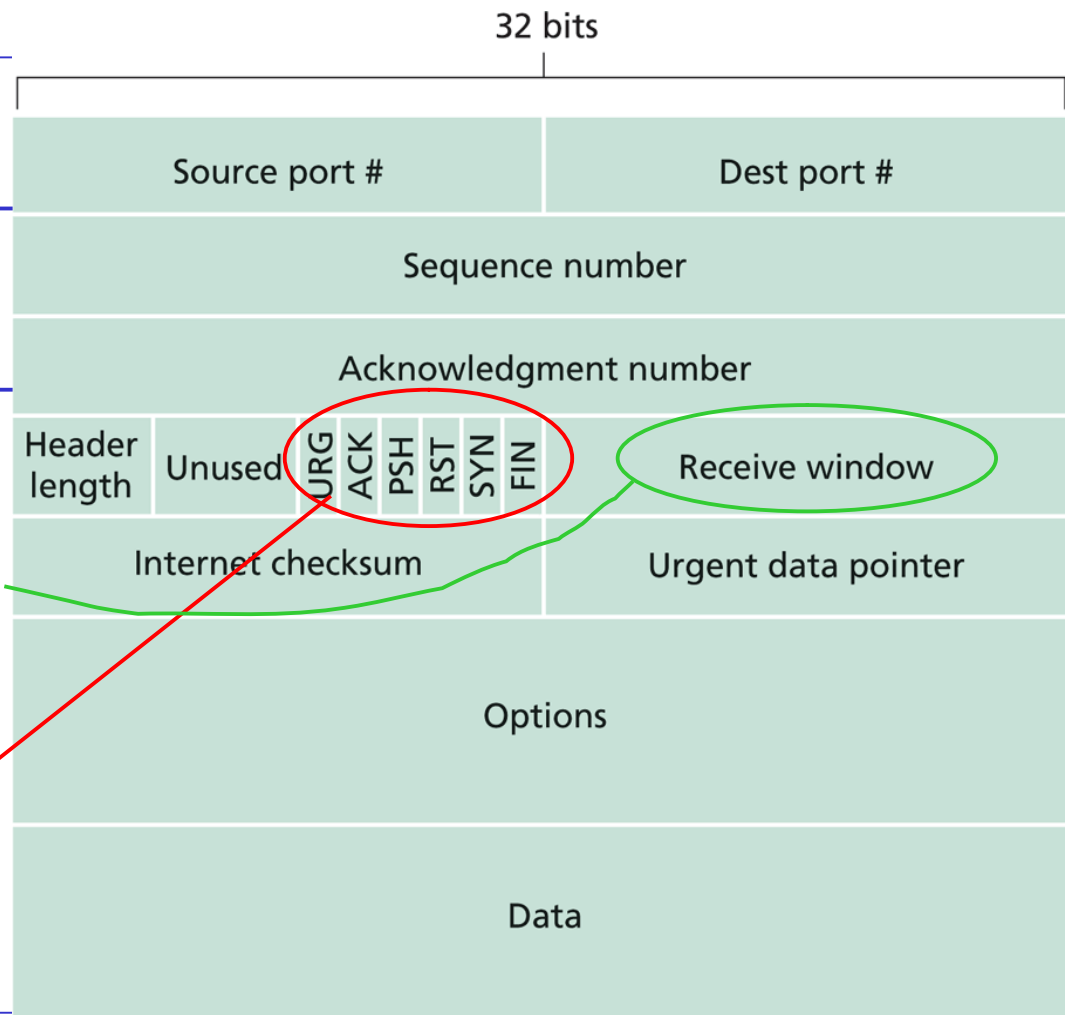
# TCP-segmentti

Otsake aina vähintään 20 B  
Optio-osa tarvittaessa

Segmentti- ja kuittaus-  
numerot **tavunumeroina**

Ikkunankoko: paljonko tilaa  
vastaanottopuskurissa (tavua)

ACK= kuittausnumero validi,  
RST (reset),  
SYN yhteydenmuodostus  
FIN yhteydenpurku  
URG, PSH ei käytetä



**Figure 3.29** ♦ TCP segment structure



# Tavunumerointi

Kuittauksia ei kuitata ja ne eivät siirrä numerointia!

Niissä ei siirretä tavuvirtaa.

Tavuvirtaa ...

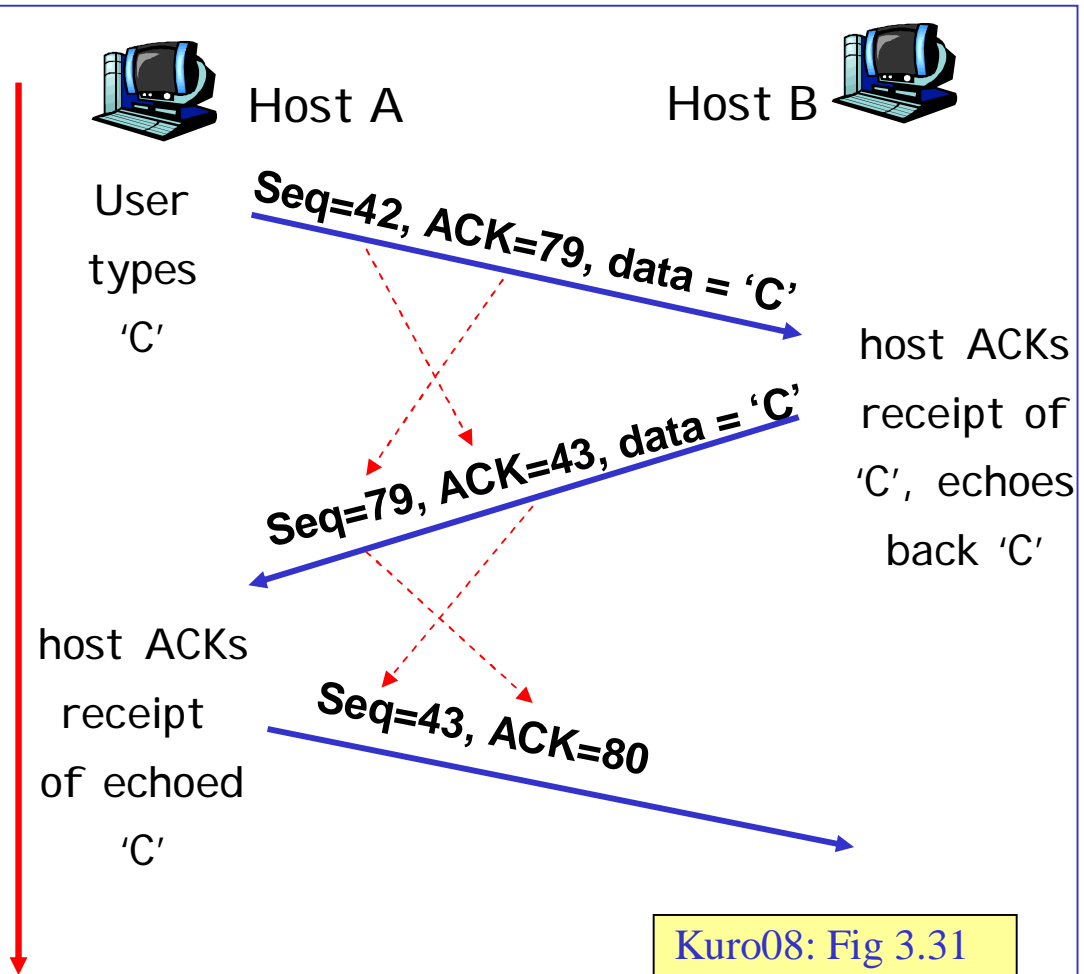
Segmentit voivat olla erikokoisia ( $\leq$  MSS)

Segmentin 'numero' =

- ensimmäisen tavun numero
- alkuarvot sovitaan yhteyttä muodostettaessa

Kuittaus

- seuraavaksi odotetun tavun numero
- kumulatiivinen
- kylkiäisenä (piggybacked) mikäli mahdollista



Kuro08: Fig 3.31

simple telnet scenario



```
NextSeqNum = InitialSeqNum; SendBase = InitialSeqNum
```

```
loop (forever) {  
  switch(event)
```

Vuonvalvonta ja/tai ruuhkanhallinta voi estää lähettämisen!

**event: data received from application above**

create TCP segment with sequence number NextSeqNum

if (timer currently not running) start timer

pass segment to IP

NextSeqNum = NextSeqNum + length(data)

Riittääkö 1 segmentti?

**event: timer timeout**

retransmit not-yet-acknowledged segment with smallest sequence number

start timer

Ajastimen arvo?

Yksi vai monta?

**event: ACK received, with ACK field value of y**

```
if (y > SendBase) {
```

```
  SendBase = y
```

```
  if (there are currently not-yet-acknowledged segments)
```

```
    start timer
```

```
}
```

Toistokuittaukset?

```
} /* end of loop forever */
```

Kuro08: Fig 3.33

## TCP: lähetys (simplified)

Comment:

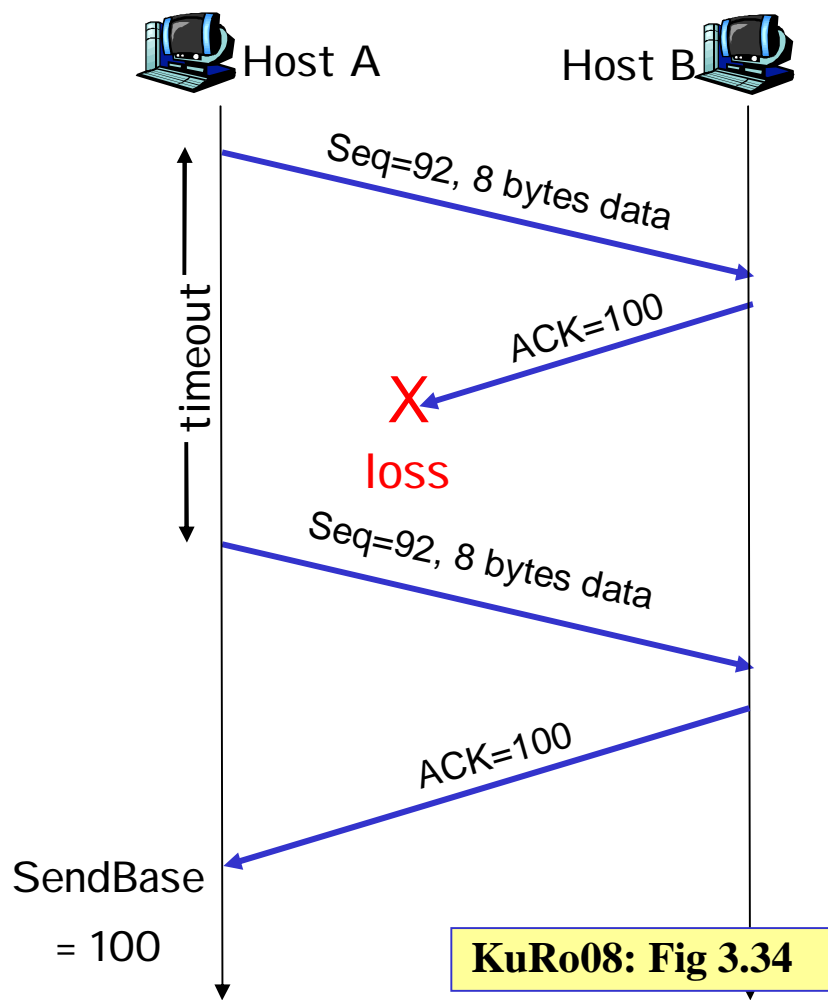
- SendBase-1: last cumulatively ack'ed byte

Example:

- SendBase-1 = 71;  
y = 73, so the rcvr wants 73+ ;  
y > SendBase, so that new data is acked

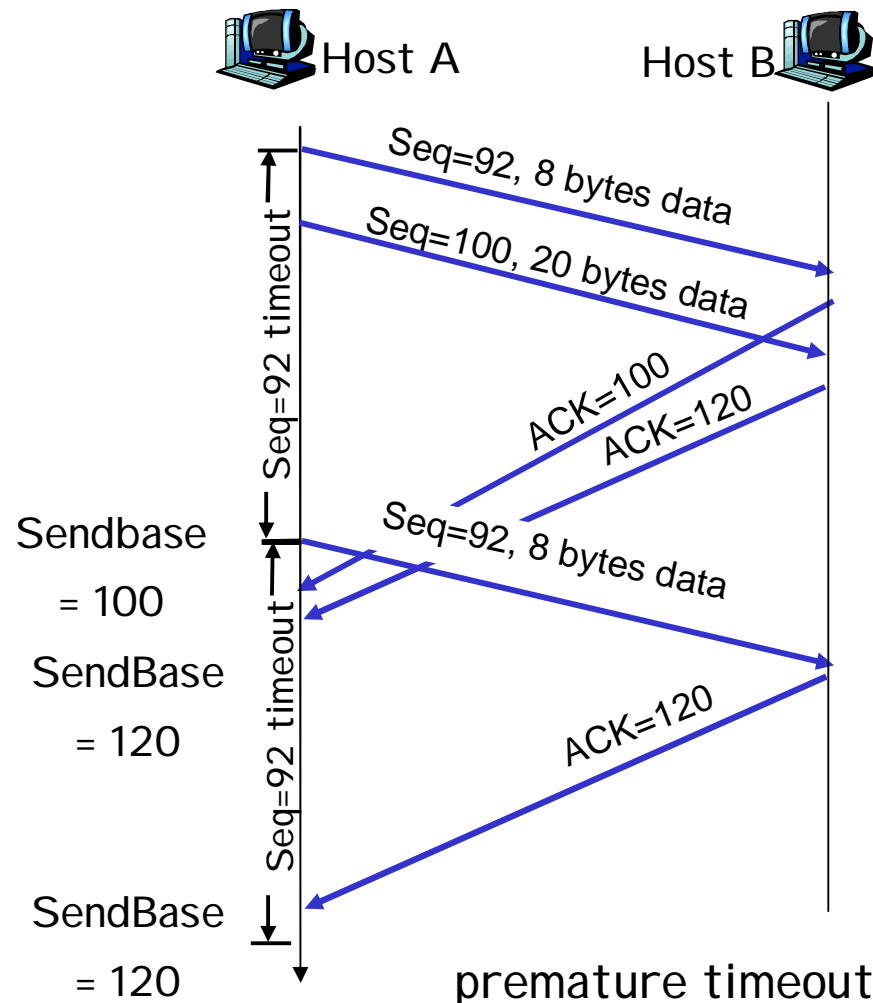


# TCP: uudelleenlähetytys



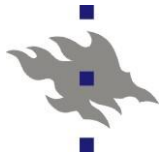
time

lost ACK scenario

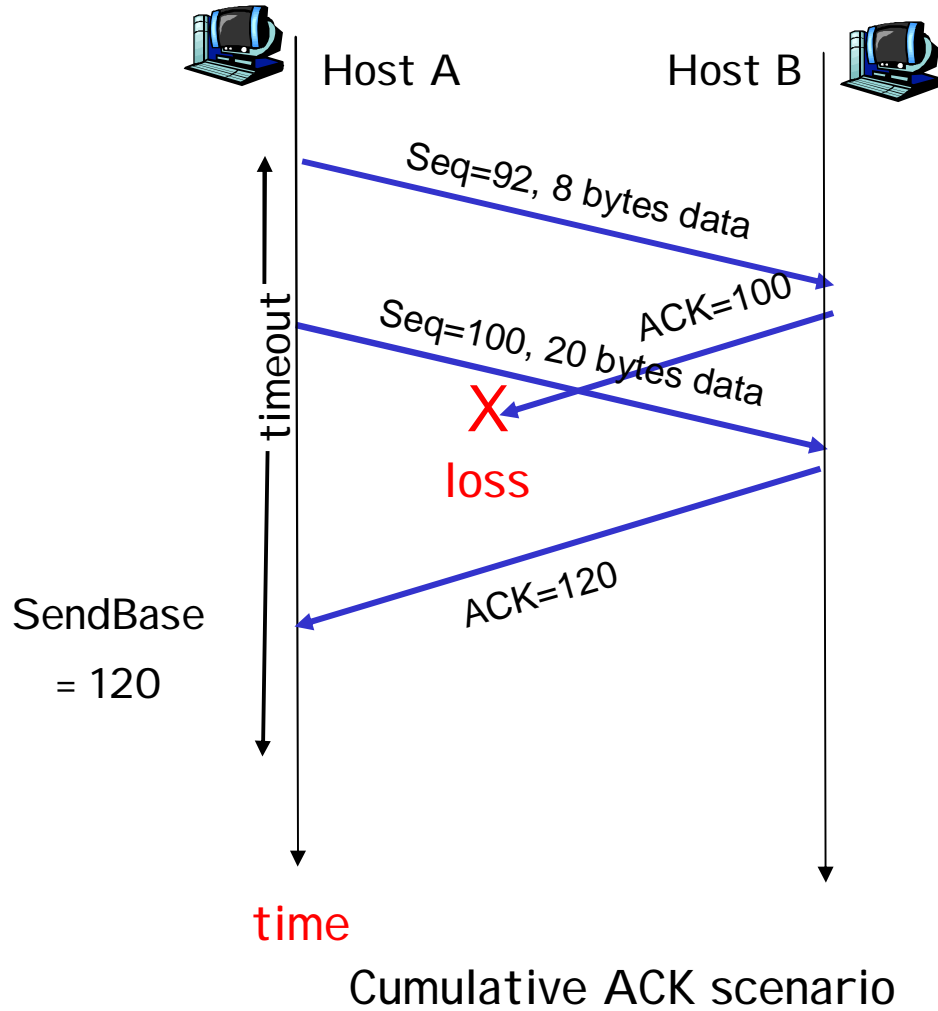


time

premature timeout



## TCP: uudelleenlähetys (2)



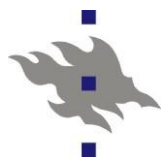
KuRo08: Fig 3.36



## TCP ACK generation [RFC 1122, RFC 2581]

Event at Receiver	TCP Receiver action
Arrival of in-order segment with expected seq #. All data up to expected seq # already ACKed	Delayed ACK. Wait up to 500ms for next segment. If no next segment, send ACK
Arrival of in-order segment with expected seq #. One other segment has ACK pending	Immediately send single cumulative ACK, ACKing both in-order segments Joka toinen kuitattava!
Arrival of out-of-order segment higher-than-expected seq. # . Gap detected	Immediately send <i>duplicate ACK</i> , indicating seq. # of next expected byte
Arrival of segment that partially or completely fills gap	Immediate send ACK, provided that segment starts at lower end of gap

TCP:n kuitattava vähintään joka toinen segmentti ja kuitausta saa viivytystä korkeintaan 500 ms.



## Nopea uudelleenlähetyks (fast retransmit)

n Timeout suhteellisen pitkä

n => aika iso viive ennen uudelleenlähetystä

n Vastaanottaja ilmoittaa puuttuvasta segmentistä toistokuittauksilla (duplicate ACK)

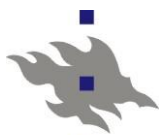
n Liukuhihnoituksen vuoksi useita segmenttejä voi olla kuittaamatta

n Jos välistä puuttuu segmentti, seurauksena on useita ACK-kuittauksia

n Jos lähettäjä saa 3 samaa segmenttiä kuittaavaa **toistokuittauksia**, se olettaa, että seuraava segmentti on kadonnut (siis kaikkiaan 4 samaa)

n Ja lähettää puuttuvan segmentin heti

n Nopea uudelleenlähetyks = **lähetä uudelleen** jo ennen kuin ajastin laukeaa eli **kolmen tuplakuittauksen jälkeen**.



## Nopea uudelleenlähetys

Sender

**event: ACK received, with ACK field value of y**

```
if (y > SendBase) { /* uuden segmentin kuittaus */
    SendBase = y
    if (there are currently not-yet-acknowledged segments)
        start timer
}
else { /* toistokuittaus */
    increment count of dup ACKs received for y
    if (count of dup ACKs received for y = 3)
    resend segment with sequence number y
}
```

a duplicate ACK for  
already ACKed segment

Nopea uudelleenlähetys  
(fast retransmit)



## Paluu n:ään vai valikoiva toisto?

### n Kumpaa TCP käyttää?

- n Liukuvan ikkunan protokolla
- n Hybridi, tavallaan 'best-of-both'

### n Go-Back-N-tyyppinen

- n Virheellisiä tai väärässä järjestyksessä tulleita **ei hyväksytä, mutta ne yleensä talletetaan puskuriin**
  - Kumulatiivinen ACK
  - Kaikkia virheellisestä lähtien ei tarvitse lähettää uudestaan

### n Valikoiva toisto

- n Kumulatiivinen ACK ei kelpaa
- n SACK-kuittaus (selective ACK), joka kertoo, mitkä segmentit vastaanotettu (RFC 2018), ei yleisesti käytössä



# Vuonvalvonta

- n Jotta lähettäjä ei tukahduta vastaanottoa

- n Siirtonopeus sovitettava vastaanottavan sovelluksen mukaan

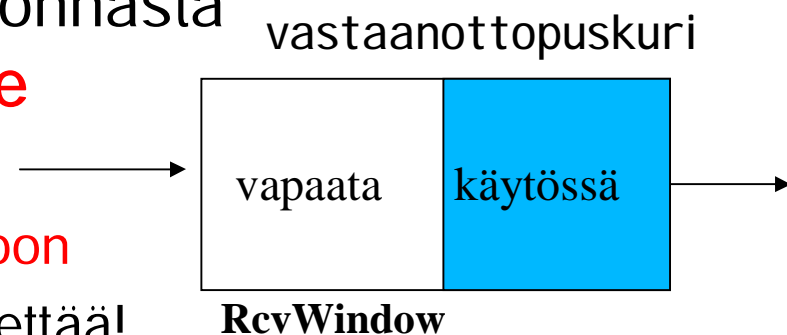
- n Kuittaus on irroitettu vuonvalvonnasta

- n **Liukuva ikkuna, koko vaihtelee**

- n Kuittaus siirtää ikkunaa

- n **Vuonvalvonta määrää ikkunankoon**

- n Kun ikkunankoko = 0, ei saa lähettää!



- n Vastaanottaja kertoo, montako tavua puskuureihin vielä mahtuu

- n TCP-segmentin otsakkeen kenttä **Receive window**

- n Sovellus lukee tavut silloin kun haluaa

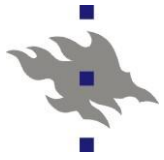
- n Koko on mukana jokaisessa TCP-segmentissä (molempiin suuntiin)

- n **Myös ruuhkanhallinta rajoittaa lähettämistä**



# Vuonvalvonta

- n Kun ikkunankoko ==0, milloin voi taas lähettää?
  
- n Jatka **lähettämällä tavun kokoisia segmenttejä**
  - n Kysely
- n Kuittaus antaa ajantasalla olevan tiedon vastaanottajan puskuritulasta
  - n Edellisen ACK:n toistokuittaus => ei tilaa
  - n Normaali ACK, kun vapaata vähintään täydelle TCP-segmentille
  
- n Miksi lähettäjä ei vain odota, että vastaanottaja kertoo, kun tilaa jälleen tulee?
  - n Entä, jos tämä kuittaus katoaa!
  - n Lähettäjä odottaa turhaan ja vastaanottaja luulee, ettei ole lähetettävää => lukkiutuminen!



# Yhteyden muodostus

## n Kolmivaiheinen kättely (three-way handshake)

- n 3 segmenttiä: SYN – SYNACK – ACK
- n Otsakkeen bittikentät
- n Viimeinen voi sisältää dataa (piggybacked)
- n Jos porttiin ei liity prosessia, vastaukseksi RST-segmentti eli yhteyttä ei voida muodostaa

## n Varaa puskuritilaa

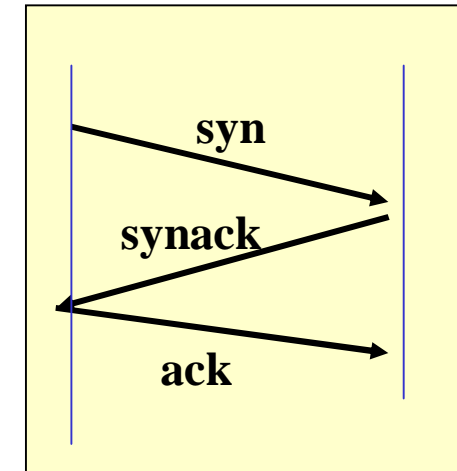
- n Lähettäjä puskuroid uudelleenlähetystä varten
- n Vastaanottaja saatujen pakettien järjestämistä varten

## n Sovitaan tavunumeroinnin alkuarvoista

- n Kuittauksia varten

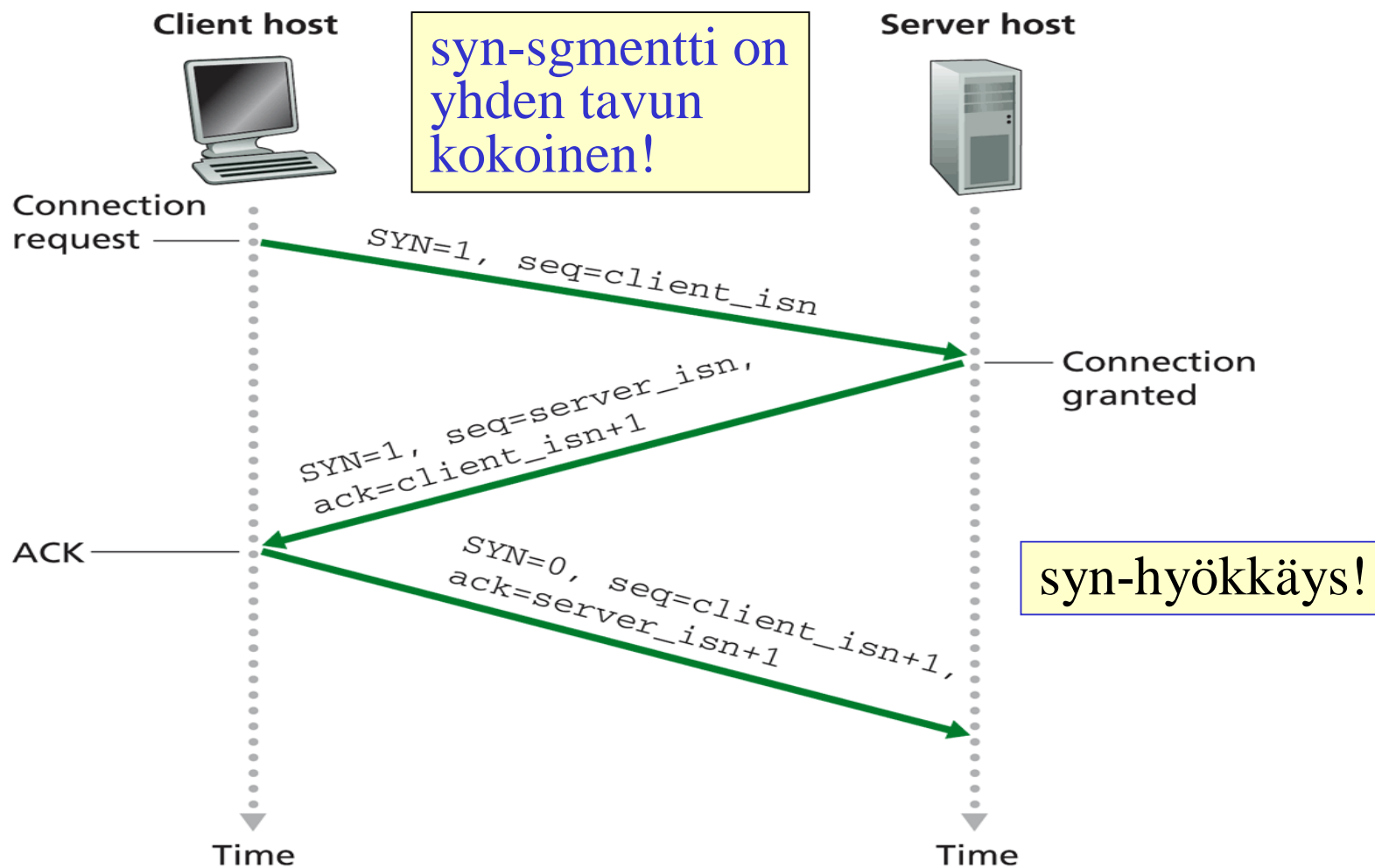
## n Ilmoita oma vastaanottoikkunan koko

- n Vuonvalvontaa varten

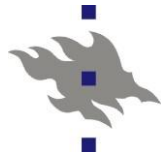




# Yhteyden muodostus



**Figure 3.39** ♦ TCP three-way handshake: segment exchange

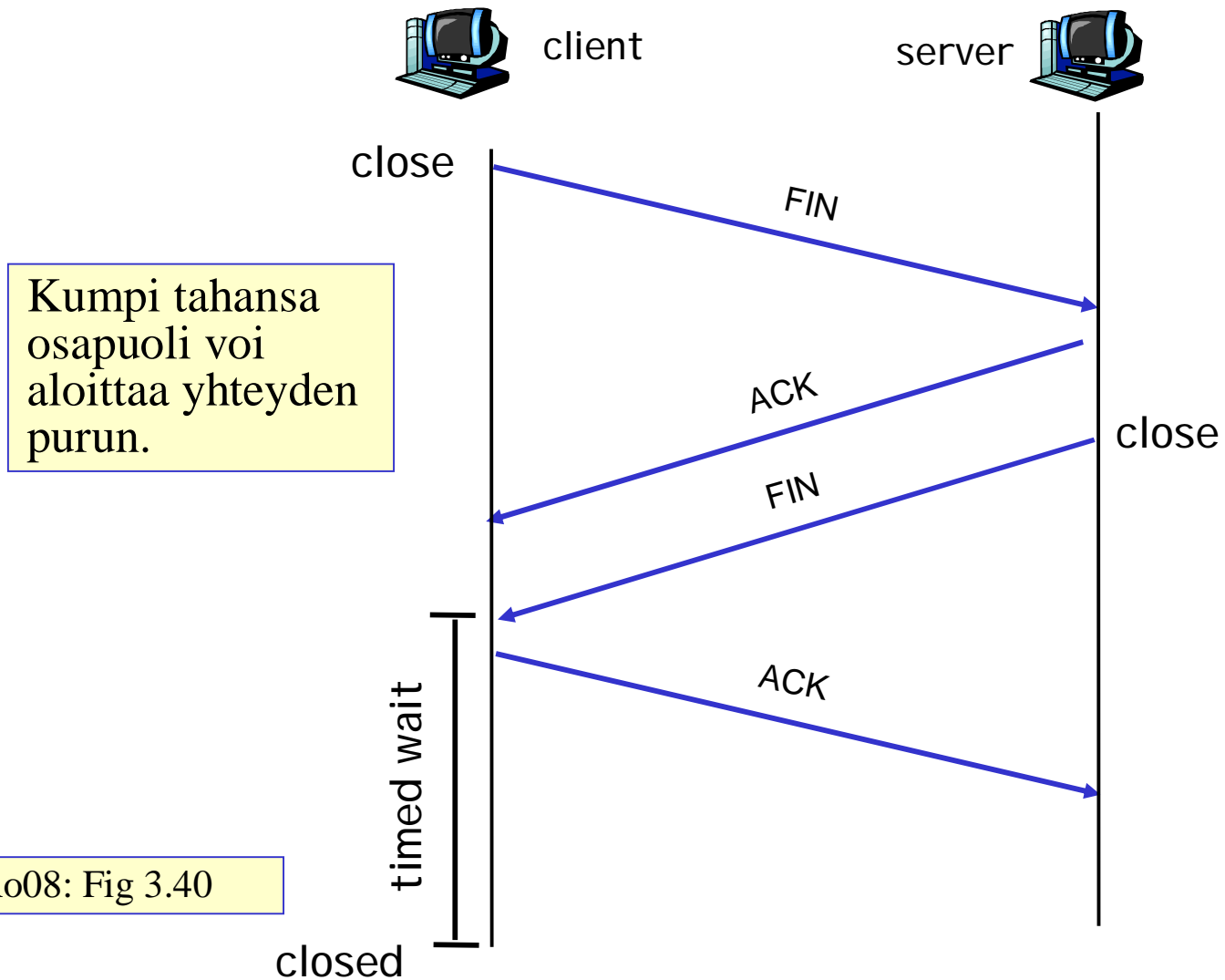


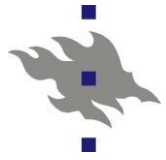
# Yhteyden purku

- n Molemmat suunnat puretaan erikseen
  - n 4 segmenttiä: FIN – ACK, FIN – ACK
- n Yhteys on kokonaan purettu, kun molemmat suunnat purettu
- n Purku käyttää ajastimia
  - n Joidenkin erikoistilanteiden hallintaan
  - n Noin  $2 \cdot$  paketin maksimaalinen elinikä
  - n Tyypillisesti 30, 60 tai 120 s

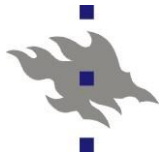


# Yhteyden purku



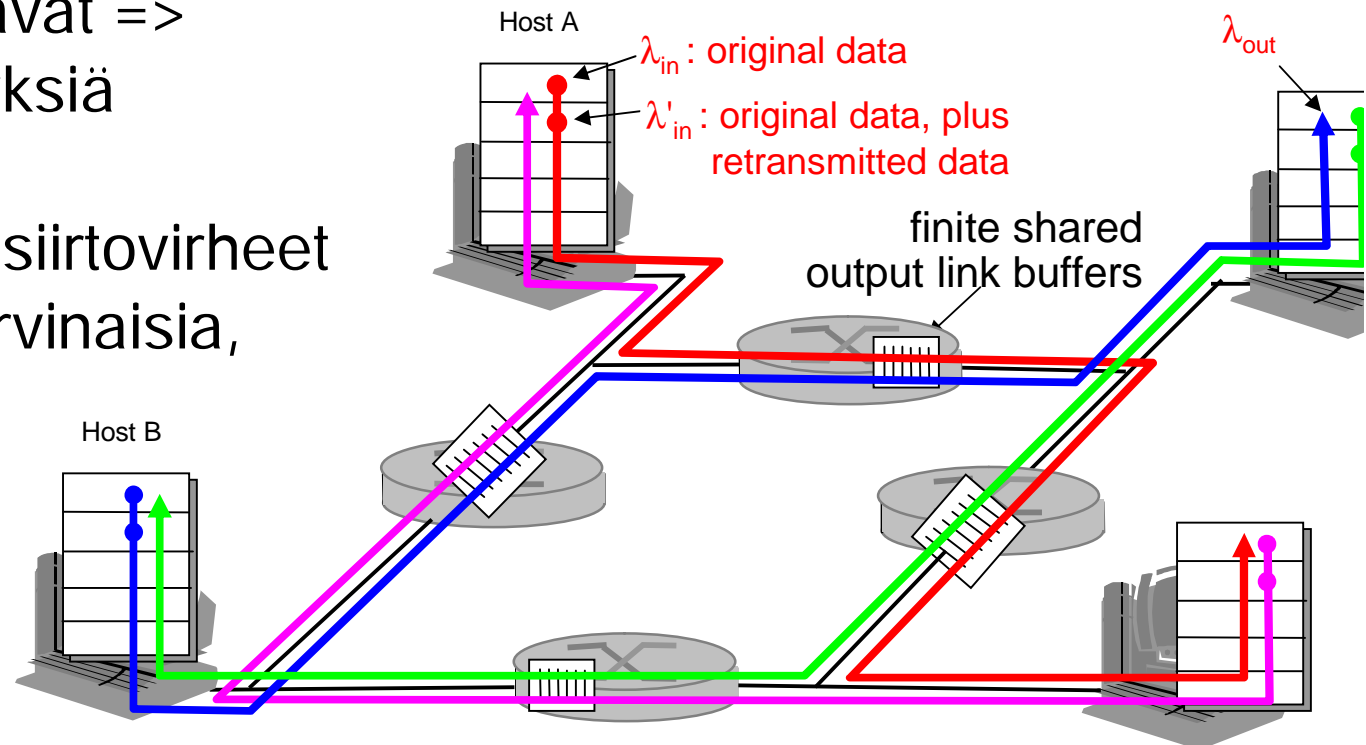


# Ruuhkanhallinta TCP:ssä



# Ruuhka

- n Reitittimelle tulee paketteja nopeammin kuin se ehtii välittää niitä eteenpäin
  - n Pitkiä jonotusviipeitä
  - n Pakettien hävittämistä (puskuritila loppuu)
- n Ajastimet laukeavat => uudelleenlähetystä
- n Lankaverkoissa siirtovirheet ovat nykyisin harvinaisia, paketin katoamisen syynä lähes aina ruuhka







## Miten ratkaista ongelma?

- n Lähettäjän on hidastettava vauhtia
- n Verkkoavusteinen ruuhkanvalvonta (network assisted congestion control)
  - n Reitittimet antavat tietoa ruuhkasta
  - n Lisäbitit kertomassa ruuhkasta tai kenttä, joka ilmoittaa yhdeydelle sallitun lähetysnopeuden (explicit rate)
- n **Päästä-päähän ruuhkanvalvonta** (end-to-end congestion control)
  - n Reitittimet eivät kerro ruuhkaantumisestaan isäntäkoneille
  - n Isäntäkoneet huomaavat itse ruuhkan lisääntyneestä pakettien katoamisesta ja uudelleenlähetyksistä
  - n TCP käyttää tätä
- n Tällä kurssilla käsitellään vain **TCP:n ruuhkanhallintaa**
  - n **Lähinnä TCP Reno -algoritmi**
  - n Ruuhkanhallintaan kehitetty runsaasti erilaisia algoritmeja



## Ruuhkaikkuna (congestion window)

- n Internetin hetkellinen kuormitus vaihtelee
- n Ruuhkaikkuna
  - n Paljonko lähettäjä saa tietyllä hetkellä kuormittaa verkkoa
  - n Paljonko lähettäjällä saa olla kuittaamattomia segmenttejä
- n Lähettäjän pääteltävä itse sopiva ikkunankoko
  - n Jos uudelleenlähetyksajastin laukeaa, on ruuhkaa
  - n Jos kuittaukset tulevat tasaisesti, ei ole ruuhkaa
- n Dynaaminen ruuhkaikkunan koko
  - n Kasvata ikkunaa ensin nopeasti, kunnes törmätään ruuhkaan
  - n Pienennä sitten ikkunaa reilusti ja kasvata varovasti
- n Lähetysikkunan raja voi tulla vastaan ensin
  - n Kuittamatta saa olla **min(lähetysikkuna, ruuhkaikkuna)**

# TCP Reno: Hidas aloitus (slow start) ja ruuhkanvälttely (congestion avoidance)

Aluksi ruuhkaikkuna = yksi segmentti

Alussa hidas siirtonopeus =  $MSS/RTT$

Kukin kuittaus kasvattaa yhdellä ruuhkaikkunan kokoa

hidas aloitus

Eksponentiaalinen kasvu

Ikkuna kaksinkertaistuu yhden RTT:n aikana

Jos uudelleenlähetys, ruuhkaikkunan kooksi 1 segmentti

Multiplicative decrease

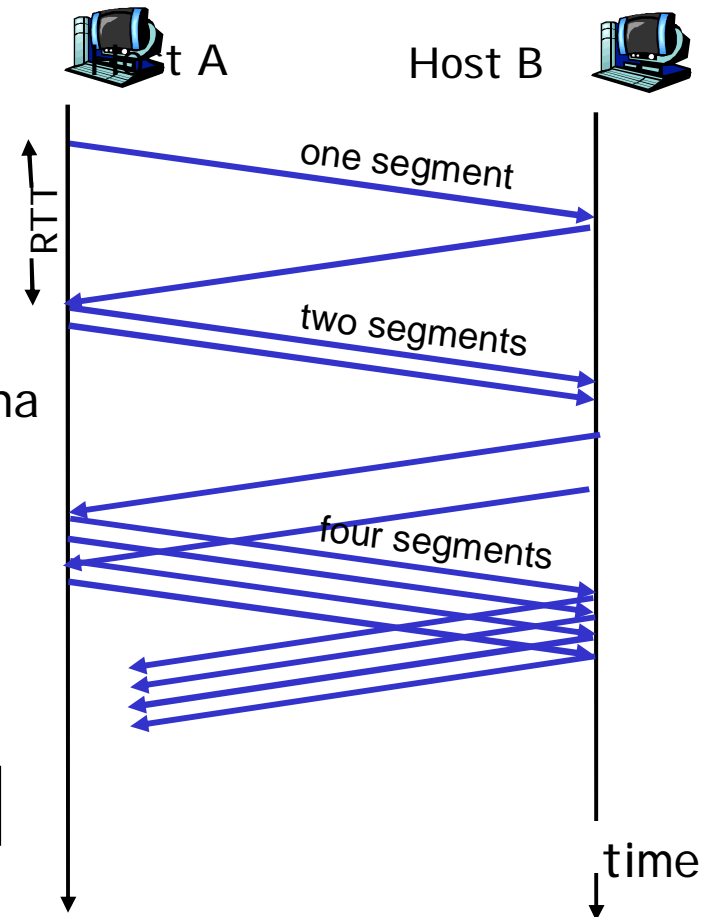
Sen jälkeen kasvata ikkunaa yksi segmentti/RTT

ruuhkanvälttely

Lineaarinen kasvu (Additive increase)

Ruuhkan välttely (congestion avoidance)

Siirtonopeus =  $CongWin / RTT$  tavua/sek





## Kynnysarvo (threshold)

n = ~ **varoitus**: tästä lähtien syytä varoa ruuhkaa

n Aluksi threshold = 64 KB

n Ajastimen laukeamisen (timeout) ja jälkeen

n Threshold = CongWin / 2

n Kynnysarvoon saakka ikkunan kasvatus eksponentilaalista

n Yhtä kuitattua segmenttiä kohden saa lähettää kaksi uutta

n Eli kaksinkertaistuu yhden RTT:n aikana

n Sen jälkeen ikkuna kasvaa lineaarisesti

ruuhkanvälttely

n Kasvaa yhdellä yhden RTT:n aikana

n Miksi näin?

# TCP Reno: Tarkennus

- n Saatu 3 ACK-kaksoiskuittausta (double ACK) (4 samaa kuittausta!)
  - n Verkkko pystyy välittämään dataa!
  - n Ei siis (pahaa) ruuhkaa, ehkä paketissa bittivirhe tai paketti kadonnut jostain muusta syystä
- n Nopea uudelleenlähetys (fast retransmit)
- n Nopea toipuminen (fast recovery)
  - n Puolita ruuhkaikkunan koko ja kasvata sitten lineaarisesti

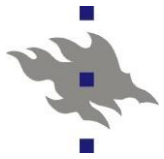
## n Timeout (= uusi hidas aloitus)

- n Verkkko pahasti ruuhkautunut!
- n Pudota ikkunankoko arvoon 1
- n Kasvata eksponentiaalisesti kynnyсарvoon asti
- n Kasvata sitten lineaarisesti

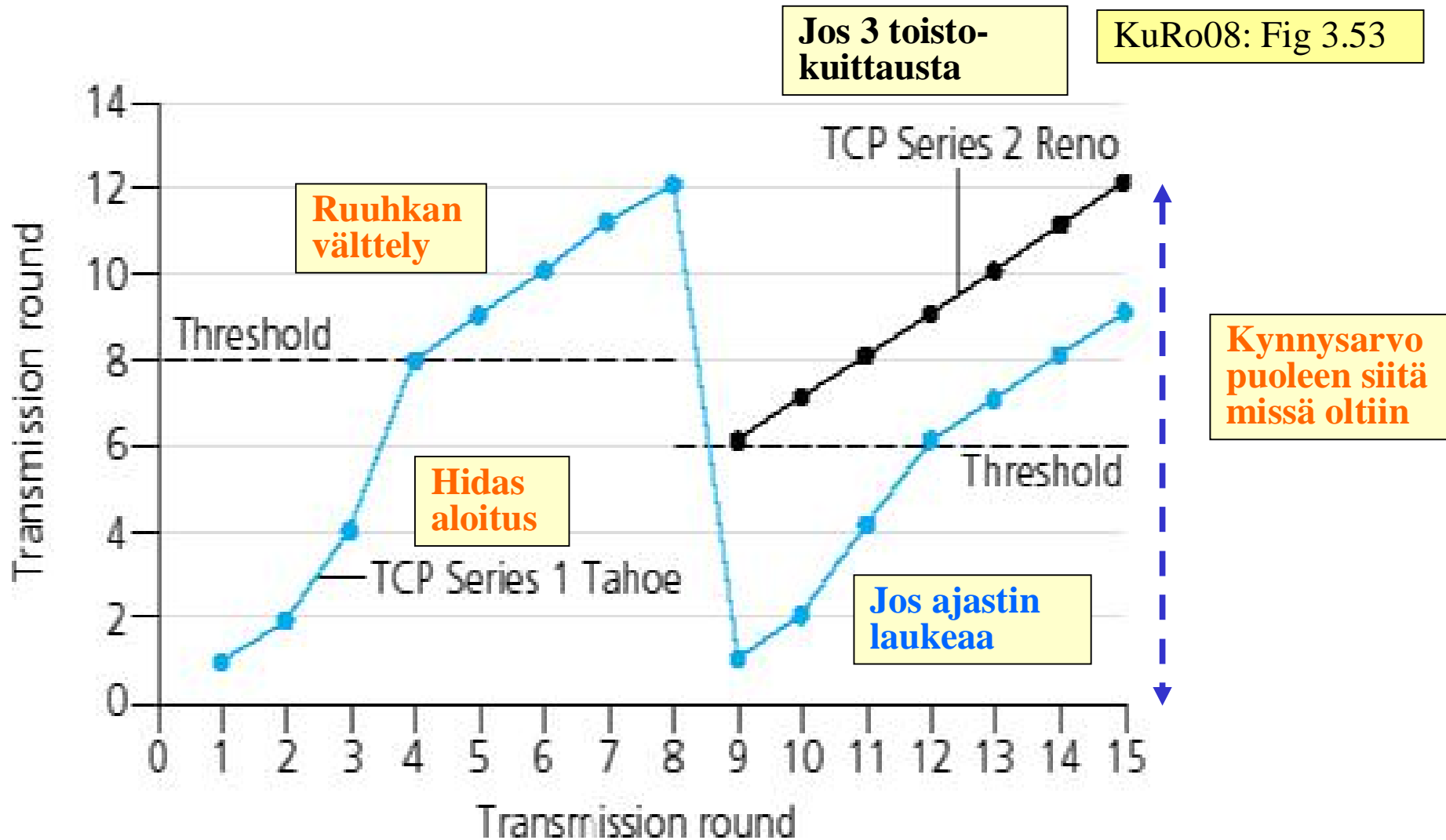
Hidas aloitus

ruuhkanvälttely

Vanha TCP Tahoe  
pudotti aina kokoon 1.



# TCP Tahoe vs. TCP Reno





# TCP Reno Ruuhkanhallinta

State	Event	TCP Sender Action	Commentary
Slow Start (SS)	ACK receipt for previously unacked data	CongWin = CongWin + MSS, If (CongWin > Threshold) set state to "Congestion Avoidance"	Resulting in a doubling of CongWin every RTT
Congestion Avoidance (CA)	ACK receipt for previously unacked data	CongWin = CongWin + MSS * (MSS/CongWin)	Additive increase, resulting in increase of CongWin by 1 MSS every RTT
SS or CA	Loss event detected by triple duplicate ACK	Threshold = CongWin/2, CongWin = Threshold, Set state to "Congestion Avoidance"	Fast recovery, implementing multiplicative decrease. CongWin will not drop below 1 MSS.
SS or CA	Timeout	Threshold = CongWin/2, CongWin = 1 MSS, Set state to "Slow Start"	Enter slow start
SS or CA	Duplicate ACK	Increment duplicate ACK count for segment being acked	CongWin and Threshold not changed

KuRo08: Table 3.3

## ■ ■ ■ Ajastimen arvo?

- n Aseta ajastin, kun segmentti on lähetetty
- n Liian lyhyt aika
  - n Ennenaikainen timeout, turha uudelleenlähetyks
  - n Turhat ruuhkatoiminnot
- n Liian pitkä aika
  - n Turhan hidas reagointi segmentin katoamiseen
  - n Ei huomata ruuhkaa ajoissa
- n Alkujaan: *Timeoutinterval = 2\*RTT*
- n *Kuittausaika vaihtelee suuresti ja nopeasti =>käytössä dynaaminen arvo*
  - n *Saadaan jatkuvien mittausten perusteella*
- n *Jos ajastin laukeaa, tuplaa Timeout*
  - n *Exponential backoff*



# Ajastimen arvo?

n Timeoutinterval = EstimatedRTT + 4 DevRTT

n Arvioi kiertoviive eli EstimatedRTT

n Mittaa jokaisen lähetetyn segmentin kiertoviive tai noin RTT:n välein normaalien lähetysten aikana.

n Laske painotettu arvo, tyypillisesti  $\alpha = 1/8 = 0.125$

$$\text{EstimatedRTT} = (1-\alpha) * \text{EstimatesRTT} + \alpha * \text{SampleRTT}$$

n Huomioi poikkeama

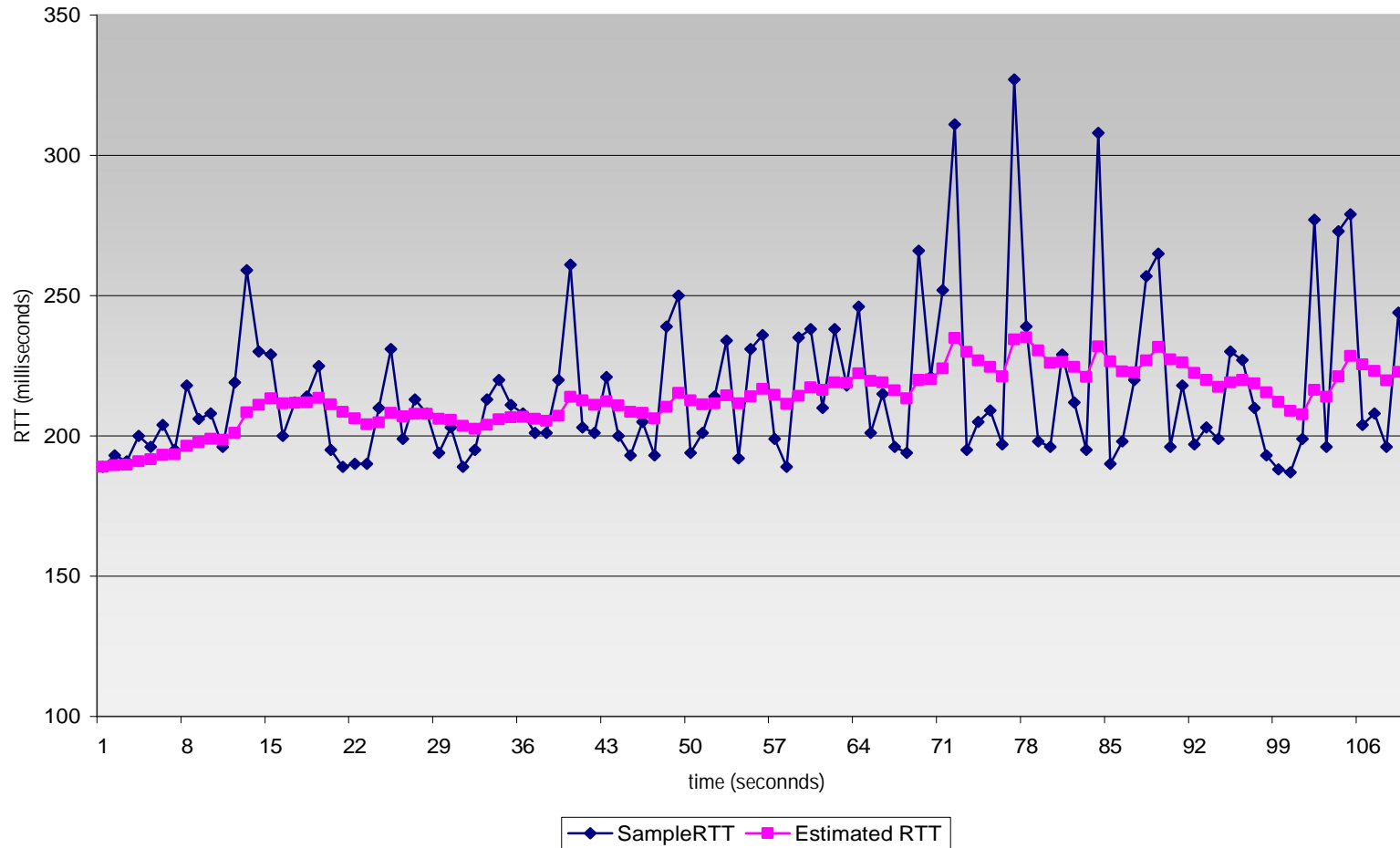
tyypillisesti  $\beta=0.25$

$$\text{DevRTT} = (1-\beta) * \text{DevRTT} + \beta * |\text{SampleRTT} - \text{EstimatedRTT}|$$

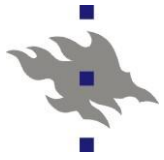


# Esimerkki

RTT: gaia.cs.umass.edu to fantasia.eurecom.fr



**Kuro05:Fig 3.32 RTT samples and RTT estimates**



# Ajastinajan kaksinkertaistaminen

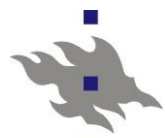
- n Aina kun saadaan lähetettäväksi sovellukselta dataa tai saadaan kuittaus jo lähetettyyn segmenttiin, otetaan käyttöön tuorein estimoitu ajastimen arvo.
- n **Ajastimen laukeaminen =>**
  - n Kun segmentti lähetetään uudelleen, kaksinkertaistetaan ajastimen arvo.
  - n Jos sama segmentti joudutaan lähettämään useaan kertaan uudestaan, niin ajastimen arvo aina kaksinkertaistetaan.

**Esimerkki:** Lähetetään segmentti 100 ja ajastimen arvo 2.5 sekuntia.

Ajastin laukeaa ja lähetetään segmentti 100 uudestaan ja nyt ajastimen arvo on 5 sekuntia.

Kun kuittausta ei tule 5 sekunnin sisällä, niin ajastin taas laukea ja segmentti 100 lähetetään vielä kerran. Nyt ajastimen arvoksi asetetaan 10 sekuntia.

Tähän saadaan kuittaus ajoissa ja siirrytään lähettämään segmenttiä 1100. Ajastimen arvoksi asetetaan tuorein estimoitu arvo 3.2 sekuntia.



## Onko TCP reilu?

### Kohdellaanko TCP-yhteyksiä reilusti?

n Jokainen reitittimessä kulkeva yhteys kärsii ruuhkasta

n Vain TCP-yhteydet kiltisti vähentävät lähetystään

n AIMD: additive increase, multiplicative decrease

n Sovellus voi avata monta rinnakkaista yhteyttä

n Onko tämä reilua muita kohtaan?

n UDP?

n Ei ruuhkanhallintaa, ei välitä ruuhkasta eikä vähennä lähetystä

n Multimediasovellukset käyttävät UDP:tä, työntävät dataa vakionopeudella ja sietävät katoamisia

n TCP-ystävällinen reititin?



## Kertauskysymyksiä

- n TCP vs. UDP?
- n Miten tehdä kuljetuspalvelusta luotettava?
- n Keskeisimmät TCP:n otsakkeessa olevat tiedot?
- n Mitä tapahtuu TCP:n yhteydenmuodostuksessa?
- n Vuonvalvonta?
- n Ruuhkanhallinta?

ks. Kurssikirja s. 293

