

COMP 445
Data Communications & Computer networks
Winter 2022

Transport Layer

- ✓ Transport-Layer services
- ✓ Multiplexing and demultiplexing
- ✓ UDP
- ✓ Reliable data transfer
- ✓ TCP
- ✓ Principles of congestion control
- ✓ TCP Congestion control

Transport Layer

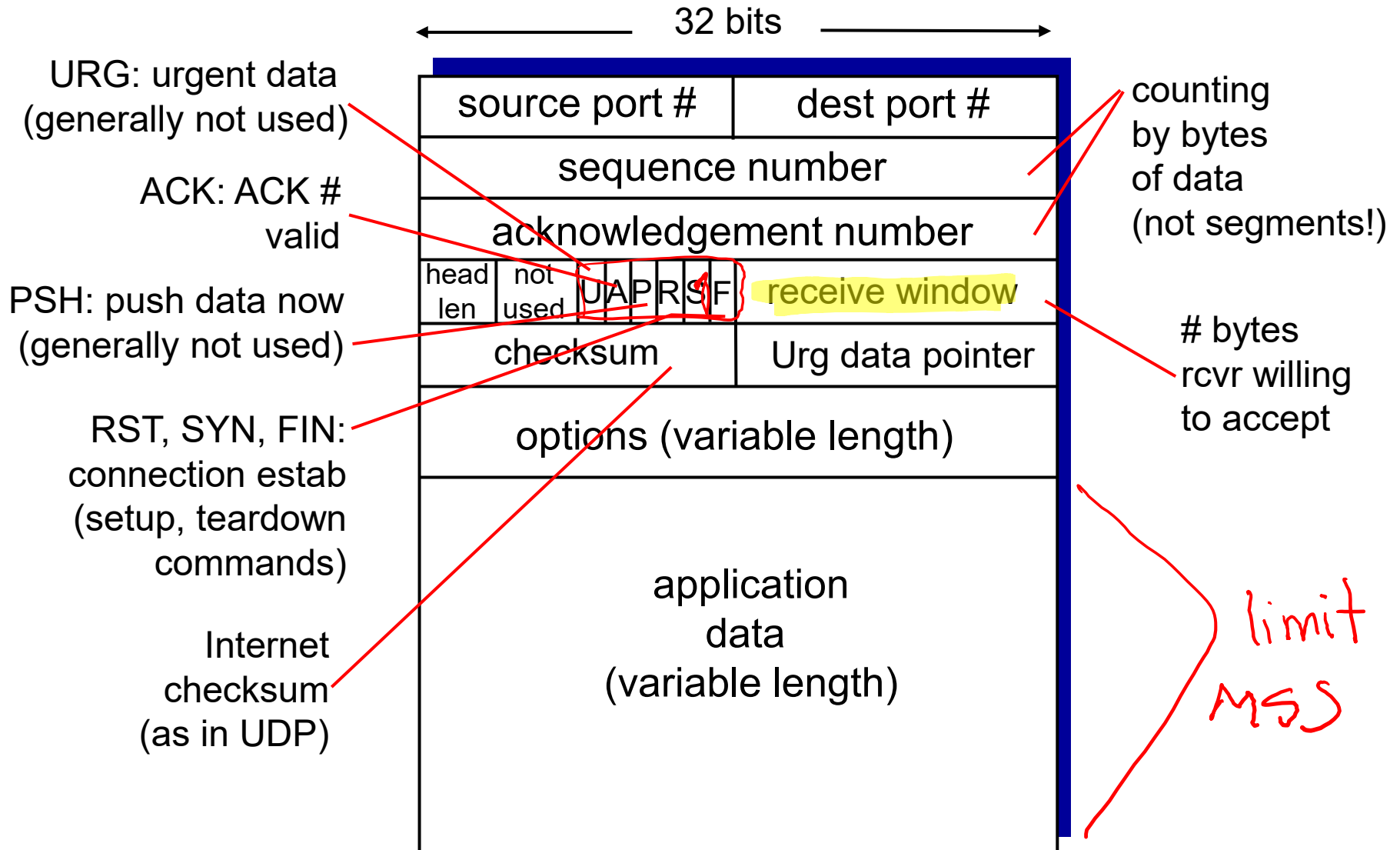
- ✓ Transport-Layer services
- ✓ Multiplexing and demultiplexing
- ✓ UDP
- ✓ Reliable data transfer
- ✓ TCP
- ✓ Principles of congestion control
- ✓ TCP Congestion control

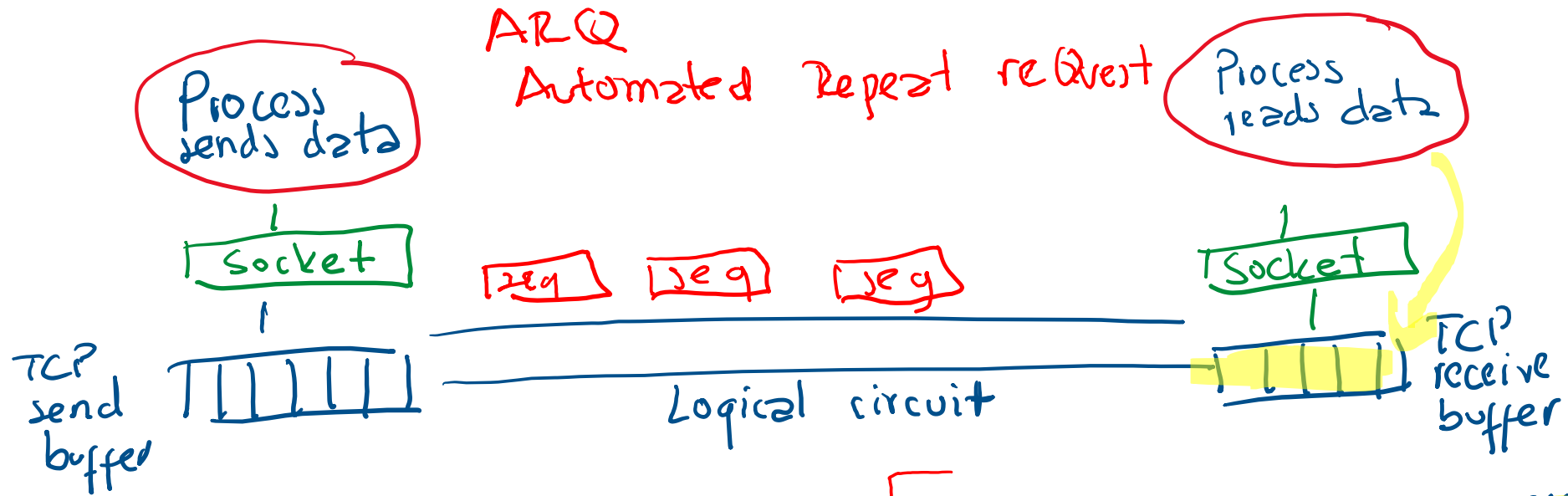
TCP: Overview

RFCs: 793, 1122, 1323, 2018, 2581

- **point-to-point:**
 - one sender, one receiver
- **reliable, in-order *byte stream*:**
 - no “message boundaries”
- **pipelined:**
 - TCP congestion and flow control set window size
- **full duplex data:**
 - bi-directional data flow in same connection
 - MSS: maximum segment size
- **connection-oriented:**
 - handshaking (exchange of control msgs) initializes sender, receiver state before data exchange
- **flow controlled:**
 - sender will not overwhelm receiver

TCP segment structure



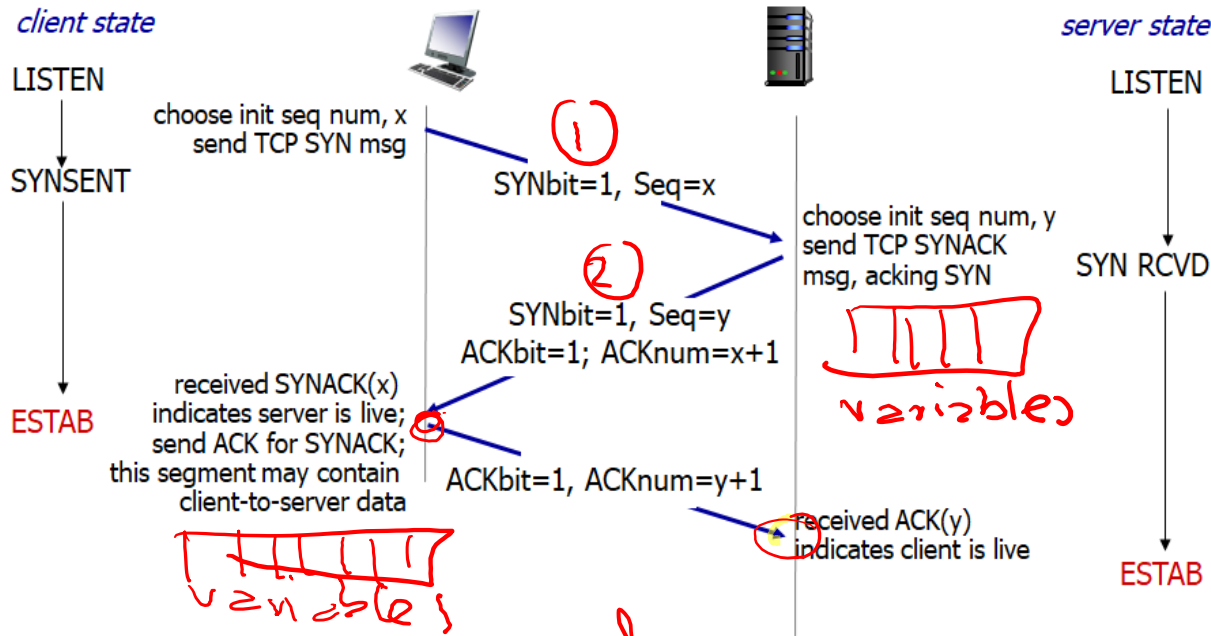
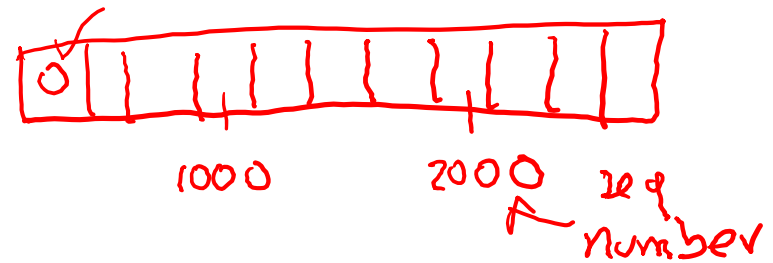


- ① When client app creates socket
 - TCP establishes session with server
 - 3-way handshaking
- ② TCP grabs data from send socket (up to MSS) and creates segment
 - starts exchange of segments

1460 bytes (typical MSS)

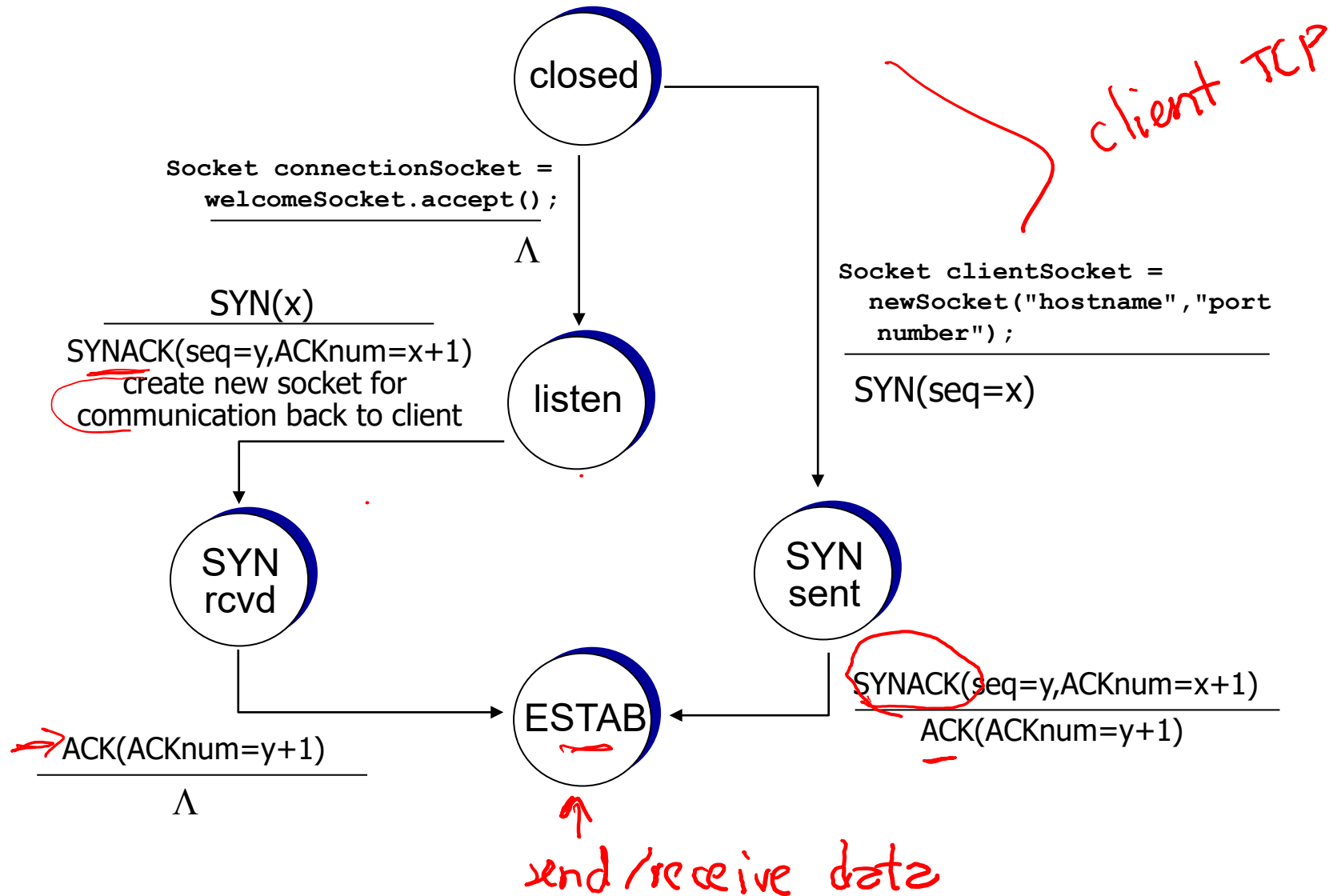
- ③ Perform RDT for error-recovery
 - Establishes RTT
 - Exchange using hybrid ARQ
 - Uses fast-retransmit
- ④ And performs flow-control
- ⑤ And performs congestion control
- ⑥ Closes the connection

TCP 3-way handshake



- ① SYN, no data, client seq number
- ② SYNACK, no data, ACK client seq num, server seq number
- ③ ACK, data (optional), ACK server seq number

TCP 3-way handshake: FSM



TCP seq. numbers, ACKs

sequence numbers:

- byte stream “**number**” of first byte in segment's data

acknowledgements:

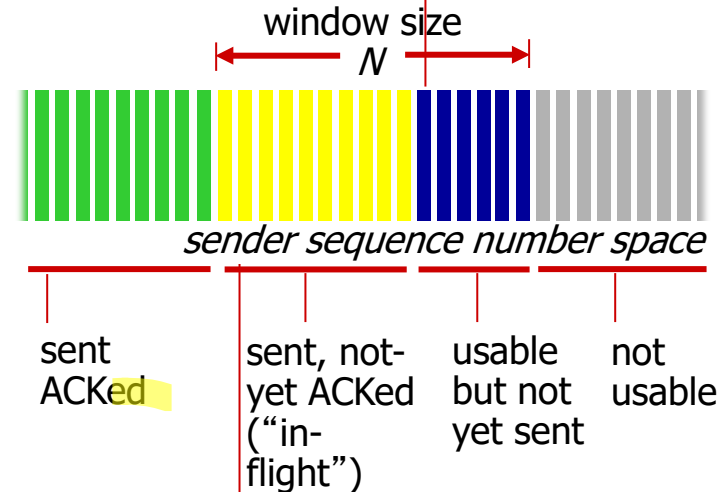
- seq # of next byte expected from other side
- cumulative ACK

Q: how receiver handles out-of-order segments

- **A:** TCP spec doesn't say,
- up to implementor

outgoing segment from sender

source port #	dest port #
sequence number	
acknowledgement number	
	rwnd
checksum	urg pointer

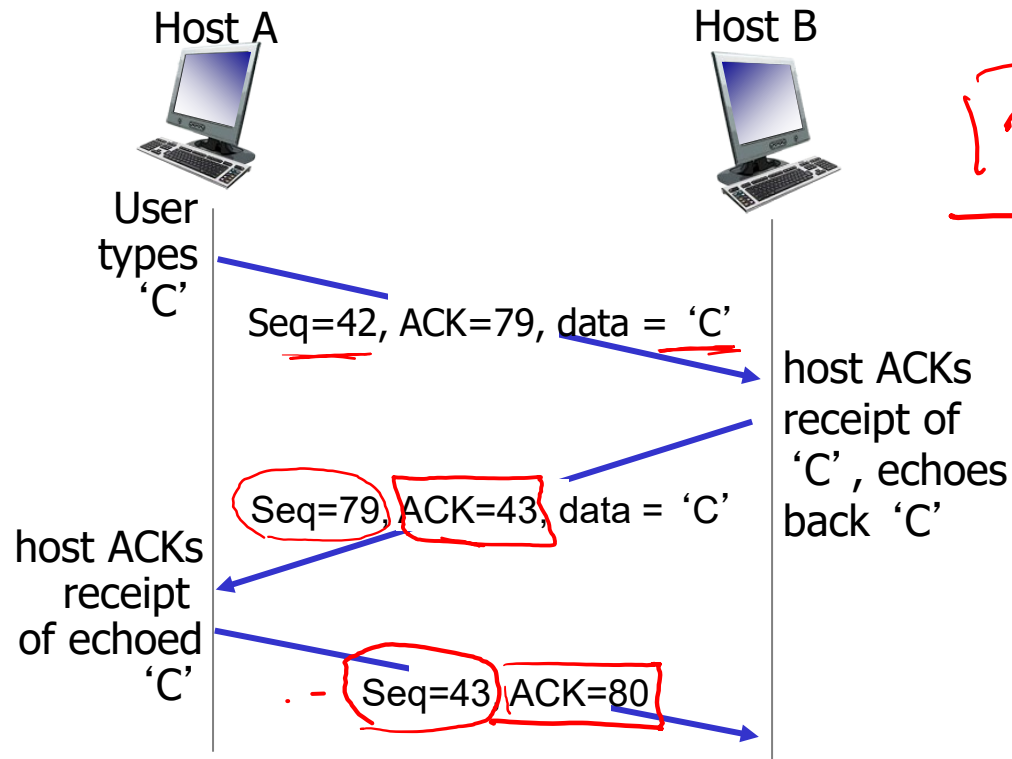


incoming segment to sender

source port #	dest port #
sequence number	
acknowledgement number	
	rwnd
checksum	urg pointer

data Transport Layer 3-9

TCP seq. numbers, ACKs



simple telnet scenario

TCP round trip time, timeout

Q: how to set TCP timeout value?

- longer than RTT
 - but RTT varies
- *too short*: premature timeout, unnecessary retransmissions
- *too long*: slow reaction to segment loss

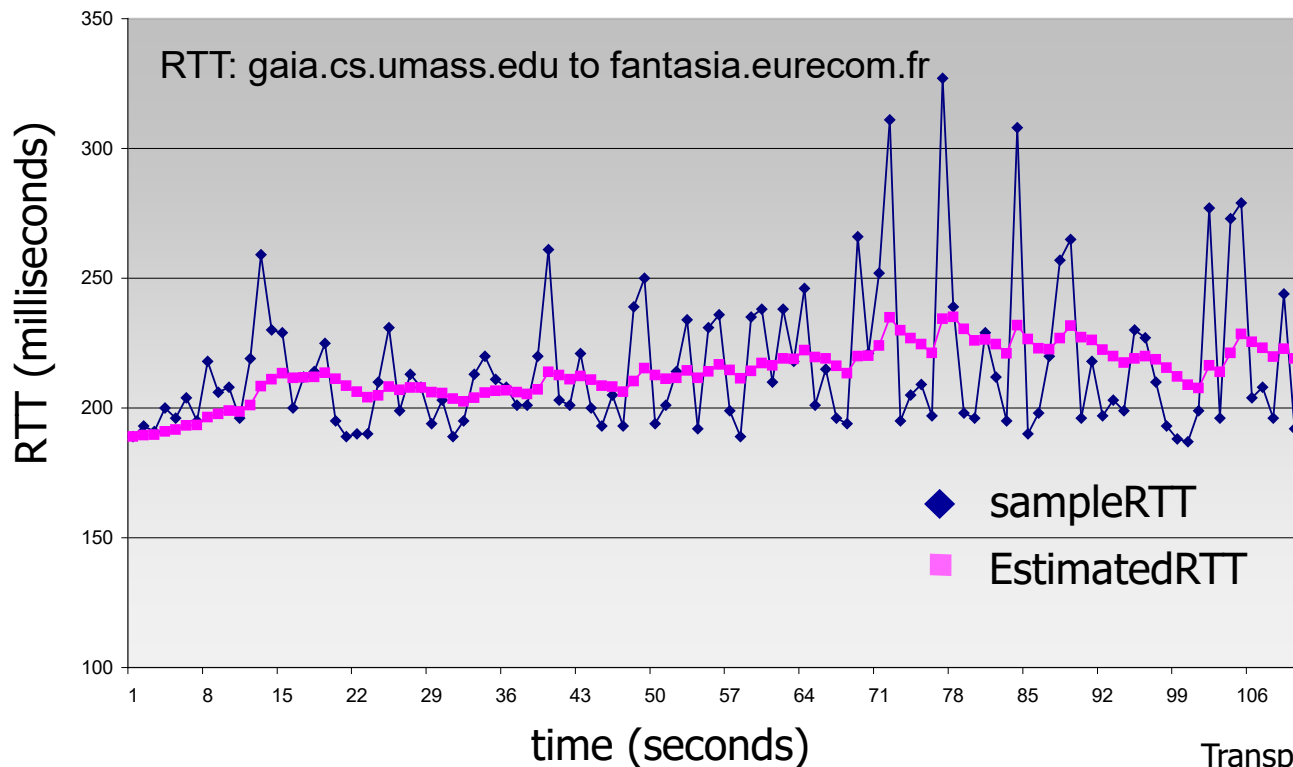
Q: how to estimate RTT?

- **SampleRTT**: measured time from segment transmission until ACK receipt
 - ignore retransmissions
- **SampleRTT** will vary, want estimated RTT “smoother”
 - average several *recent* measurements, not just current **SampleRTT**

TCP round trip time, timeout

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

- exponential weighted moving average
- influence of past sample decreases exponentially fast
- typical value: $\alpha = 0.125$



TCP round trip time, timeout

- **timeout interval:** **EstimatedRTT** plus “safety margin”
 - large variation in **EstimatedRTT** → larger safety margin
- estimate SampleRTT deviation from EstimatedRTT:

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

(typically, $\beta = 0.25$)

$$\text{TimeoutInterval} = \text{EstimatedRTT} + 4 * \text{DevRTT}$$



↑
estimated RTT

↑
“safety margin”

TCP reliable data transfer

- TCP creates rdt service on top of IP's unreliable service

- pipelined segments
- cumulative acks
- single retransmission timer

- retransmissions triggered by:

- timeout events
- duplicate acks

let's initially consider simplified TCP sender:

- ignore duplicate acks
- ignore flow control, congestion control

TCP sender events:

data rcvd from app:

- create segment with seq #
- seq # is byte-stream number of first data byte in segment
- start timer if not already running
 - think of timer as for oldest unacked segment
 - expiration interval: `TimeoutInterval`

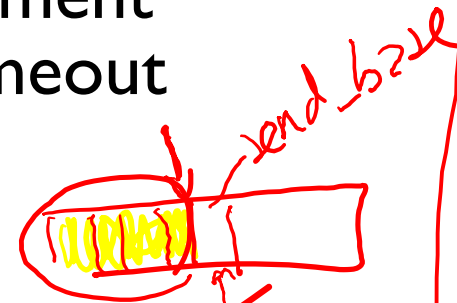
↓
estimated RTT dev RTT

timeout:

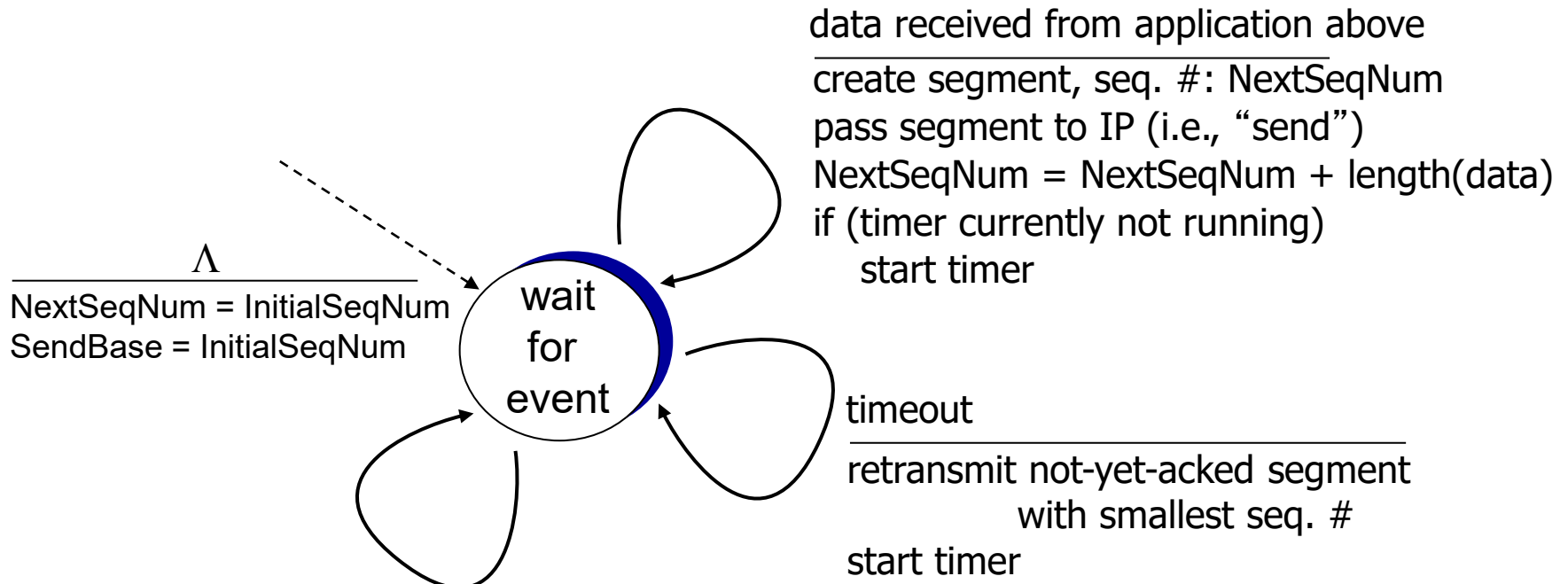
- retransmit segment that caused timeout
- restart timer

ack rcvd:

- if ack acknowledges previously unacked segments
 - update what is known to be ACKed
 - start timer if there are still unacked segments



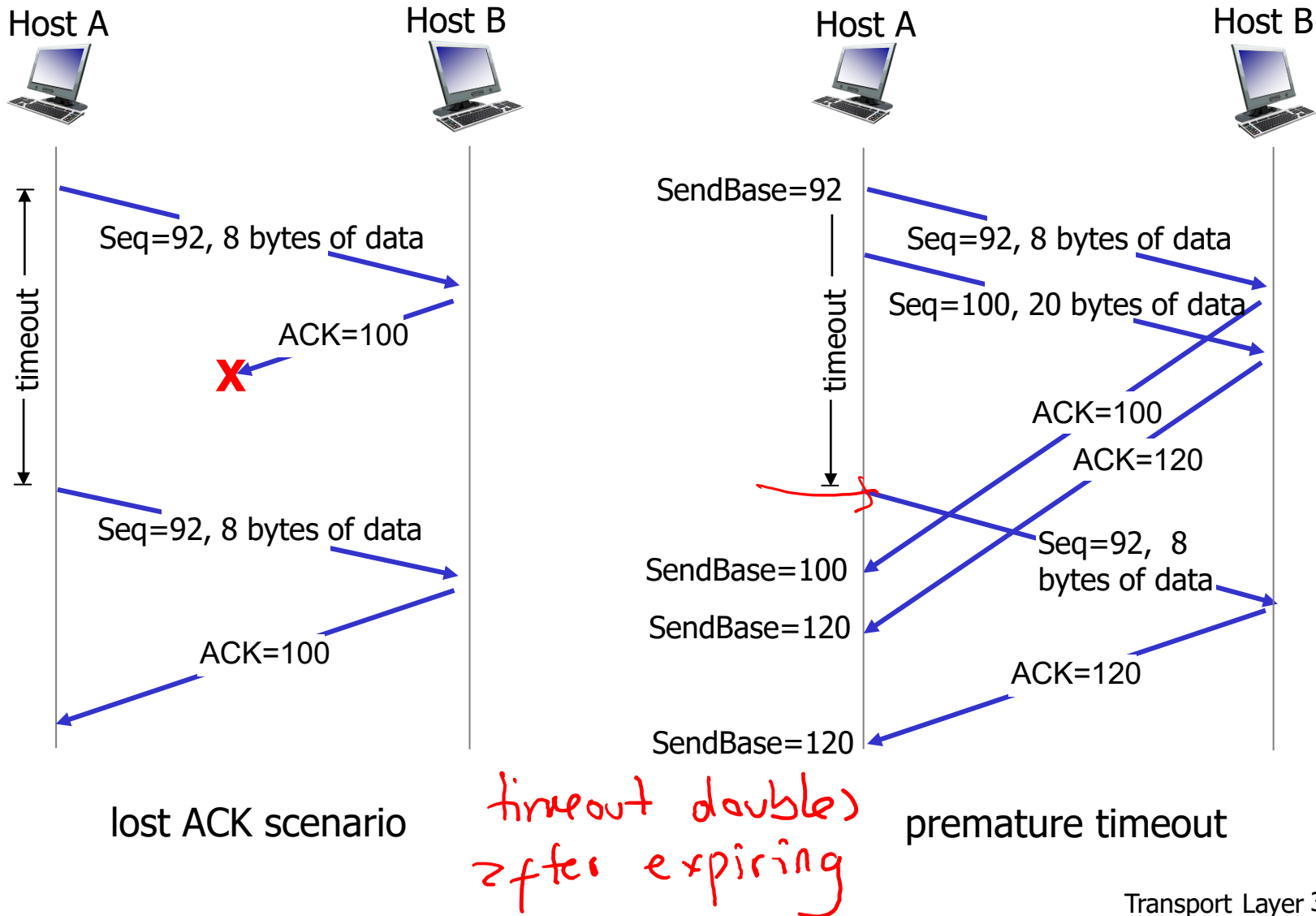
TCP sender (simplified)



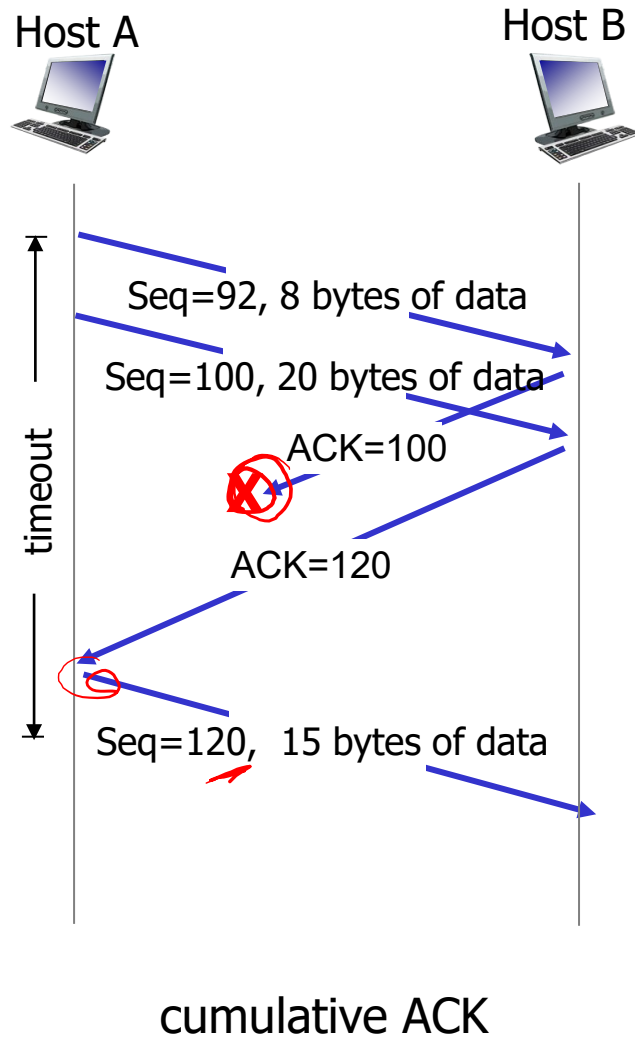
ACK received, with ACK field value y

```
if (y > SendBase) {  
    SendBase = y  
    /* SendBase-1: last cumulatively ACKed byte */  
    if (there are currently not-yet-acked segments)  
        start timer  
    else stop timer  
}
```


TCP: retransmission scenarios



TCP: retransmission scenarios



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

P P 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

arrival of out-of-order segment higher-than-expect seq. # .

immediately send duplicate ACK, indicating seq. # of next expected byte

Gap detected

~~1~~ ~~2~~ ~~3~~ ~~4~~ ~~5~~ ~~6~~ ~~7~~ ~~8~~ ~~9~~ ~~10~~ ~~11~~ ~~12~~ ~~13~~ ~~14~~ ~~15~~ ~~16~~ ~~17~~ ~~18~~ ~~19~~ ~~20~~ ~~21~~ ~~22~~ ~~23~~ ~~24~~ ~~25~~ ~~26~~ ~~27~~ ~~28~~ ~~29~~ ~~30~~ ~~31~~ ~~32~~ ~~33~~ ~~34~~ ~~35~~ ~~36~~ ~~37~~ ~~38~~ ~~39~~ ~~40~~ ~~41~~ ~~42~~ ~~43~~ ~~44~~ ~~45~~ ~~46~~ ~~47~~ ~~48~~ ~~49~~ ~~50~~ ~~51~~ ~~52~~ ~~53~~ ~~54~~ ~~55~~ ~~56~~ ~~57~~ ~~58~~ ~~59~~ ~~60~~ ~~61~~ ~~62~~ ~~63~~ ~~64~~ ~~65~~ ~~66~~ ~~67~~ ~~68~~ ~~69~~ ~~70~~ ~~71~~ ~~72~~ ~~73~~ ~~74~~ ~~75~~ ~~76~~ ~~77~~ ~~78~~ ~~79~~ ~~80~~ ~~81~~ ~~82~~ ~~83~~ ~~84~~ ~~85~~ ~~86~~ ~~87~~ ~~88~~ ~~89~~ ~~90~~ ~~91~~ ~~92~~ ~~93~~ ~~94~~ ~~95~~ ~~96~~ ~~97~~ ~~98~~ ~~99~~ ~~100~~ ~~101~~ ~~102~~ ~~103~~ ~~104~~ ~~105~~ ~~106~~ ~~107~~ ~~108~~ ~~109~~ ~~110~~ ~~111~~ ~~112~~ ~~113~~ ~~114~~ ~~115~~ ~~116~~ ~~117~~ ~~118~~ ~~119~~ ~~120~~ ~~121~~ ~~122~~ ~~123~~ ~~124~~ ~~125~~ ~~126~~ ~~127~~ ~~128~~ ~~129~~ ~~130~~ ~~131~~ ~~132~~ ~~133~~ ~~134~~ ~~135~~ ~~136~~ ~~137~~ ~~138~~ ~~139~~ ~~140~~ ~~141~~ ~~142~~ ~~143~~ ~~144~~ ~~145~~ ~~146~~ ~~147~~ ~~148~~ ~~149~~ ~~150~~ ~~151~~ ~~152~~ ~~153~~ ~~154~~ ~~155~~ ~~156~~ ~~157~~ ~~158~~ ~~159~~ ~~160~~ ~~161~~ ~~162~~ ~~163~~ ~~164~~ ~~165~~ ~~166~~ ~~167~~ ~~168~~ ~~169~~ ~~170~~ ~~171~~ ~~172~~ ~~173~~ ~~174~~ ~~175~~ ~~176~~ ~~177~~ ~~178~~ ~~179~~ ~~180~~ ~~181~~ ~~182~~ ~~183~~ ~~184~~ ~~185~~ ~~186~~ ~~187~~ ~~188~~ ~~189~~ ~~190~~ ~~191~~ ~~192~~ ~~193~~ ~~194~~ ~~195~~ ~~196~~ ~~197~~ ~~198~~ ~~199~~ ~~200~~ ~~201~~ ~~202~~ ~~203~~ ~~204~~ ~~205~~ ~~206~~ ~~207~~ ~~208~~ ~~209~~ ~~210~~ ~~211~~ ~~212~~ ~~213~~ ~~214~~ ~~215~~ ~~216~~ ~~217~~ ~~218~~ ~~219~~ ~~220~~ ~~221~~ ~~222~~ ~~223~~ ~~224~~ ~~225~~ ~~226~~ ~~227~~ ~~228~~ ~~229~~ ~~230~~ ~~231~~ ~~232~~ ~~233~~ ~~234~~ ~~235~~ ~~236~~ ~~237~~ ~~238~~ ~~239~~ ~~240~~ ~~241~~ ~~242~~ ~~243~~ ~~244~~ ~~245~~ ~~246~~ ~~247~~ ~~248~~ ~~249~~ ~~250~~ ~~251~~ ~~252~~ ~~253~~ ~~254~~ ~~255~~ ~~256~~ ~~257~~ ~~258~~ ~~259~~ ~~260~~ ~~261~~ ~~262~~ ~~263~~ ~~264~~ ~~265~~ ~~266~~ ~~267~~ ~~268~~ ~~269~~ ~~270~~ ~~271~~ ~~272~~ ~~273~~ ~~274~~ ~~275~~ ~~276~~ ~~277~~ ~~278~~ ~~279~~ ~~280~~ ~~281~~ ~~282~~ ~~283~~ ~~284~~ ~~285~~ ~~286~~ ~~287~~ ~~288~~ ~~289~~ ~~290~~ ~~291~~ ~~292~~ ~~293~~ ~~294~~ ~~295~~ ~~296~~ ~~297~~ ~~298~~ ~~299~~ ~~300~~ ~~301~~ ~~302~~ ~~303~~ ~~304~~ ~~305~~ ~~306~~ ~~307~~ ~~308~~ ~~309~~ ~~310~~ ~~311~~ ~~312~~ ~~313~~ ~~314~~ ~~315~~ ~~316~~ ~~317~~ ~~318~~ ~~319~~ ~~320~~ ~~321~~ ~~322~~ ~~323~~ ~~324~~ ~~325~~ ~~326~~ ~~327~~ ~~328~~ ~~329~~ ~~330~~ ~~331~~ ~~332~~ ~~333~~ ~~334~~ ~~335~~ ~~336~~ ~~337~~ ~~338~~ ~~339~~ ~~340~~ ~~341~~ ~~342~~ ~~343~~ ~~344~~ ~~345~~ ~~346~~ ~~347~~ ~~348~~ ~~349~~ ~~350~~ ~~351~~ ~~352~~ ~~353~~ ~~354~~ ~~355~~ ~~356~~ ~~357~~ ~~358~~ ~~359~~ ~~360~~ ~~361~~ ~~362~~ ~~363~~ ~~364~~ ~~365~~ ~~366~~ ~~367~~ ~~368~~ ~~369~~ ~~370~~ ~~371~~ ~~372~~ ~~373~~ ~~374~~ ~~375~~ ~~376~~ ~~377~~ ~~378~~ ~~379~~ ~~380~~ ~~381~~ ~~382~~ ~~383~~ ~~384~~ ~~385~~ ~~386~~ ~~387~~ ~~388~~ ~~389~~ ~~390~~ ~~391~~ ~~392~~ ~~393~~ ~~394~~ ~~395~~ ~~396~~ ~~397~~ ~~398~~ ~~399~~ ~~400~~ ~~401~~ ~~402~~ ~~403~~ ~~404~~ ~~405~~ ~~406~~ ~~407~~ ~~408~~ ~~409~~ ~~410~~ ~~411~~ ~~412~~ ~~413~~ ~~414~~ ~~415~~ ~~416~~ ~~417~~ ~~418~~ ~~419~~ ~~420~~ ~~421~~ ~~422~~ ~~423~~ ~~424~~ ~~425~~ ~~426~~ ~~427~~ ~~428~~ ~~429~~ ~~430~~ ~~431~~ ~~432~~ ~~433~~ ~~434~~ ~~435~~ ~~436~~ ~~437~~ ~~438~~ ~~439~~ ~~440~~ ~~441~~ ~~442~~ ~~443~~ ~~444~~ ~~445~~ ~~446~~ ~~447~~ ~~448~~ ~~449~~ ~~450~~ ~~451~~ ~~452~~ ~~453~~ ~~454~~ ~~455~~ ~~456~~ ~~457~~ ~~458~~ ~~459~~ ~~460~~ ~~461~~ ~~462~~ ~~463~~ ~~464~~ ~~465~~ ~~466~~ ~~467~~ ~~468~~ ~~469~~ ~~470~~ ~~471~~ ~~472~~ ~~473~~ ~~474~~ ~~475~~ ~~476~~ ~~477~~ ~~478~~ ~~479~~ ~~480~~ ~~481~~ ~~482~~ ~~483~~ ~~484~~ ~~485~~ ~~486~~ ~~487~~ ~~488~~ ~~489~~ ~~490~~ ~~491~~ ~~492~~ ~~493~~ ~~494~~ ~~495~~ ~~496~~ ~~497~~ ~~498~~ ~~499~~ ~~500~~ ~~501~~ ~~502~~ ~~503~~ ~~504~~ ~~505~~ ~~506~~ ~~507~~ ~~508~~ ~~509~~ ~~510~~ ~~511~~ ~~512~~ ~~513~~ ~~514~~ ~~515~~ ~~516~~ ~~517~~ ~~518~~ ~~519~~ ~~520~~ ~~521~~ ~~522~~ ~~523~~ ~~524~~ ~~525~~ ~~526~~ ~~527~~ ~~528~~ ~~529~~ ~~530~~ ~~531~~ ~~532~~ ~~533~~ ~~534~~ ~~535~~ ~~536~~ ~~537~~ ~~538~~ ~~539~~ ~~540~~ ~~541~~ ~~542~~ ~~543~~ ~~544~~ ~~545~~ ~~546~~ ~~547~~ ~~548~~ ~~549~~ ~~550~~ ~~551~~ ~~552~~ ~~553~~ ~~554~~ ~~555~~ ~~556~~ ~~557~~ ~~558~~ ~~559~~ ~~560~~ ~~561~~ ~~562~~ ~~563~~ ~~564~~ ~~565~~ ~~566~~ ~~567~~ ~~568~~ ~~569~~ ~~570~~ ~~571~~ ~~572~~ ~~573~~ ~~574~~ ~~575~~ ~~576~~ ~~577~~ ~~578~~ ~~579~~ ~~580~~ ~~581~~ ~~582~~ ~~583~~ ~~584~~ ~~585~~ ~~586~~ ~~587~~ ~~588~~ ~~589~~ ~~590~~ ~~591~~ ~~592~~ ~~593~~ ~~594~~ ~~595~~ ~~596~~ ~~597~~ ~~598~~ ~~599~~ ~~600~~ ~~601~~ ~~602~~ ~~603~~ ~~604~~ ~~605~~ ~~606~~ ~~607~~ ~~608~~ ~~609~~ ~~610~~ ~~611~~ ~~612~~ ~~613~~ ~~614~~ ~~615~~ ~~616~~ ~~617~~ ~~618~~ ~~619~~ ~~620~~ ~~621~~ ~~622~~ ~~623~~ ~~624~~ ~~625~~ ~~626~~ ~~627~~ ~~628~~ ~~629~~ ~~630~~ ~~631~~ ~~632~~ ~~633~~ ~~634~~ ~~635~~ ~~636~~ ~~637~~ ~~638~~ ~~639~~ ~~640~~ ~~641~~ ~~642~~ ~~643~~ ~~644~~ ~~645~~ ~~646~~ ~~647~~ ~~648~~ ~~649~~ ~~650~~ ~~651~~ ~~652~~ ~~653~~ ~~654~~ ~~655~~ ~~656~~ ~~657~~ ~~658~~ ~~659~~ ~~660~~ ~~661~~ ~~662~~ ~~663~~ ~~664~~ ~~665~~ ~~666~~ ~~667~~ ~~668~~ ~~669~~ ~~670~~ ~~671~~ ~~672~~ ~~673~~ ~~674~~ ~~675~~ ~~676~~ ~~677~~ ~~678~~ ~~679~~ ~~680~~ ~~681~~ ~~682~~ ~~683~~ ~~684~~ ~~685~~ ~~686~~ ~~687~~ ~~688~~ ~~689~~ ~~690~~ ~~691~~ ~~692~~ ~~693~~ ~~694~~ ~~695~~ ~~696~~ ~~697~~ ~~698~~ ~~699~~ ~~700~~ ~~701~~ ~~702~~ ~~703~~ ~~704~~ ~~705~~ ~~706~~ ~~707~~ ~~708~~ ~~709~~ ~~710~~ ~~711~~ ~~712~~ ~~713~~ ~~714~~ ~~715~~ ~~716~~ ~~717~~ ~~718~~ ~~719~~ ~~720~~ ~~721~~ ~~722~~ ~~723~~ ~~724~~ ~~725~~ ~~726~~ ~~727~~ ~~728~~ ~~729~~ ~~730~~ ~~731~~ ~~732~~ ~~733~~ ~~734~~ ~~735~~ ~~736~~ ~~737~~ ~~738~~ ~~739~~ ~~740~~ ~~741~~ ~~742~~ ~~743~~ ~~744~~ ~~745~~ ~~746~~ ~~747~~ ~~748~~ ~~749~~ ~~750~~ ~~751~~ ~~752~~ ~~753~~ ~~754~~ ~~755~~ ~~756~~ ~~757~~ ~~758~~ ~~759~~ ~~760~~ ~~761~~ ~~762~~ ~~763~~ ~~764~~ ~~765~~ ~~766~~ ~~767~~ ~~768~~ ~~769~~ ~~770~~ ~~771~~ ~~772~~ ~~773~~ ~~774~~ ~~775~~ ~~776~~ ~~777~~ ~~778~~ ~~779~~ ~~780~~ ~~781~~ ~~782~~ ~~783~~ ~~784~~ ~~785~~ ~~786~~ ~~787~~ ~~788~~ ~~789~~ ~~790~~ ~~791~~ ~~792~~ ~~793~~ ~~794~~ ~~795~~ ~~796~~ ~~797~~ ~~798~~ ~~799~~ ~~800~~ ~~801~~ ~~802~~ ~~803~~ ~~804~~ ~~805~~ ~~806~~ ~~807~~ ~~808~~ ~~809~~ ~~810~~ ~~811~~ ~~812~~ ~~813~~ ~~814~~ ~~815~~ ~~816~~ ~~817~~ ~~818~~ ~~819~~ ~~820~~ ~~821~~ ~~822~~ ~~823~~ ~~824~~ ~~825~~ ~~826~~ ~~827~~ ~~828~~ ~~829~~ ~~830~~ ~~831~~ ~~832~~ ~~833~~ ~~834~~ ~~835~~ ~~836~~ ~~837~~ ~~838~~ ~~839~~ ~~840~~ ~~841~~ ~~842~~ ~~843~~ ~~844~~ ~~845~~ ~~846~~ ~~847~~ ~~848~~ ~~849~~ ~~850~~ ~~851~~ ~~852~~ ~~853~~ ~~854~~ ~~855~~ ~~856~~ ~~857~~ ~~858~~ ~~859~~ ~~860~~ ~~861~~ ~~862~~ ~~863~~ ~~864~~ ~~865~~ ~~866~~ ~~867~~ ~~868~~ ~~869~~ ~~870~~ ~~871~~ ~~872~~ ~~873~~ ~~874~~ ~~875~~ ~~876~~ ~~877~~ ~~878~~ ~~879~~ ~~880~~ ~~881~~ ~~882~~ ~~883~~ ~~884~~ ~~885~~ ~~886~~ ~~887~~ ~~888~~ ~~889~~ ~~890~~ ~~891~~ ~~892~~ ~~893~~ ~~894~~ ~~895~~ ~~896~~ ~~897~~ ~~898~~ ~~899~~ ~~900~~ ~~901~~ ~~902~~ ~~903~~ ~~904~~ ~~905~~ ~~906~~ ~~907~~ ~~908~~ ~~909~~ ~~910~~ ~~911~~ ~~912~~ ~~913~~ ~~914~~ ~~915~~ ~~916~~ ~~917~~ ~~918~~ ~~919~~ ~~920~~ ~~921~~ ~~922~~ ~~923~~ ~~924~~ ~~925~~ ~~926~~ ~~927~~ ~~928~~ ~~929~~ ~~930~~ ~~931~~ ~~932~~ ~~933~~ ~~934~~ ~~935~~ ~~936~~ ~~937~~ ~~938~~ ~~939~~ ~~940~~ ~~941~~ ~~942~~ ~~943~~ ~~944~~ ~~945~~ ~~946~~ ~~947~~ ~~948~~ ~~949~~ ~~950~~ ~~951~~ ~~952~~ ~~953~~ ~~954~~ ~~955~~ ~~956~~ ~~957~~ ~~958~~ ~~959~~ ~~960~~ ~~961~~ ~~962~~ ~~963~~ ~~964~~ ~~965~~ ~~966~~ ~~967~~ ~~968~~ ~~969~~ ~~970~~ ~~971~~ ~~972~~ ~~973~~ ~~974~~ ~~975~~ ~~976~~ ~~977~~ ~~978~~ ~~979~~ ~~980~~ ~~981~~ ~~982~~ ~~983~~ ~~984~~ ~~985~~ ~~986~~ ~~987~~ ~~988~~ ~~989~~ ~~990~~ ~~991~~ ~~992~~ ~~993~~ ~~994~~ ~~995~~ ~~996~~ ~~997~~ ~~998~~ ~~999~~ ~~1000~~ ~~1001~~ ~~1002~~ ~~1003~~ ~~1004~~ ~~1005~~ ~~1006~~ ~~1007~~ ~~1008~~ ~~1009~~ ~~1010~~ ~~1011~~ ~~1012~~ ~~1013~~ ~~1014~~ ~~1015~~ ~~1016~~ ~~1017~~ ~~1018~~ ~~1019~~ ~~1020~~ ~~1021~~ ~~1022~~ ~~1023~~ ~~1024~~ ~~1025~~ ~~1026~~ ~~1027~~ ~~1028~~ ~~1029~~ ~~1030~~ ~~1031~~ ~~1032~~ ~~1033~~ ~~1034~~ ~~1035~~ ~~1036~~ ~~1037~~ ~~1038~~ ~~1039~~ ~~1040~~ ~~1041~~ ~~1042~~ ~~1043~~ ~~1044~~ ~~1045~~ ~~1046~~ ~~1047~~ ~~1048~~ ~~1049~~ ~~1050~~ ~~1051~~ ~~1052~~ ~~1053~~ ~~1054~~ ~~1055~~ ~~1056~~ ~~1057~~ ~~1058~~ ~~1059~~ ~~1060~~ ~~1061~~ ~~1062~~ ~~1063~~ ~~1064~~ ~~1065~~ ~~1066~~ ~~1067~~ ~~1068~~ ~~1069~~ ~~1070~~ ~~1071~~ ~~1072~~ ~~1073~~ ~~1074~~ ~~1075~~ ~~1076~~ ~~1077~~ ~~1078~~ ~~1079~~ ~~1080~~ ~~1081~~ ~~1082~~ ~~1083~~ ~~1084~~ ~~1085~~ ~~1086~~ ~~1087~~ ~~1088~~ ~~1089~~ ~~1090~~ ~~1091~~ ~~1092~~ ~~1093~~ ~~1094~~ ~~1095~~ ~~1096~~ ~~1097~~ ~~1098~~ ~~1099~~ ~~1100~~ ~~1101~~ ~~1102~~ ~~1103~~ ~~1104~~ ~~1105~~ ~~1106~~ ~~1107~~ ~~1108~~ ~~1109~~ ~~1110~~ ~~1111~~ ~~1112~~ ~~1113~~ ~~1114~~ ~~1115~~ ~~1116~~ ~~1117~~ ~~1118~~ ~~1119~~ ~~1120~~ ~~1121~~ ~~1122~~ ~~1123~~ ~~1124~~ ~~1125~~ ~~1126~~ ~~1127~~ ~~1128~~ ~~1129~~ ~~1130~~ ~~1131~~ ~~1132~~ ~~1133~~ ~~1134~~ ~~1135~~ ~~1136~~ ~~1137~~ ~~1138~~ ~~1139~~ ~~1140~~ ~~1141~~ ~~1142~~ ~~1143~~ ~~1144~~ ~~1145~~ ~~1146~~ ~~1147~~ ~~1148~~ ~~1149~~ ~~1150~~ ~~1151~~ ~~1152~~ ~~1153~~ ~~1154~~ ~~1155~~ ~~1156~~ ~~1157~~ ~~1158~~ ~~1159~~ ~~1160~~ ~~1161~~ ~~1162~~ ~~1163~~ ~~1164~~ ~~1165~~ ~~1166~~ ~~1167~~ ~~1168~~ ~~1169~~ ~~1170~~ ~~1171~~ ~~1172~~ ~~1173~~ ~~1174~~ ~~1175~~ ~~1176~~ ~~1177~~ ~~1178~~ ~~1179~~ ~~1180~~ ~~1181~~ ~~1182~~ ~~1183~~ ~~1184~~ ~~1185~~ ~~1186~~ ~~1187~~ ~~1188~~ ~~1189~~ ~~1190~~ ~~1191~~ ~~1192~~ ~~1193~~ ~~1194~~ ~~1195~~ ~~1196~~ ~~1197~~ ~~1198~~ ~~1199~~ ~~1200~~ ~~1201~~ ~~1202~~ ~~1203~~ ~~1204~~ ~~1205~~ ~~1206~~ ~~1207~~ ~~1208~~ ~~1209~~ ~~1210~~ ~~1211~~ ~~1212~~ ~~1213~~ ~~1214~~ ~~1215~~ ~~1216~~ ~~1217~~ ~~1218~~ ~~1219~~ ~~1220~~ ~~1221~~ ~~1222~~ ~~1223~~ ~~1224~~ ~~1225~~ ~~1226~~ ~~1227~~ ~~1228~~ ~~1229~~ ~~1230~~ ~~1231~~ ~~1232~~ ~~1233~~ ~~1234~~ ~~1235~~ ~~1236~~ ~~1237~~ ~~1238~~ ~~1239~~ ~~1240~~ ~~1241~~ ~~1242~~ ~~1243~~ ~~1244~~ ~~1245~~ ~~1246~~ ~~1247~~ ~~1248~~ ~~1249~~ ~~1250~~ ~~1251~~ ~~1252~~ ~~1253~~ ~~1254~~ ~~1255~~ ~~1256~~ ~~1257~~ ~~1258~~ ~~1259~~ ~~1260~~ ~~1261~~ ~~1262~~ ~~1263~~ ~~1264~~ ~~1265~~ ~~1266~~ ~~1267~~ ~~1268~~ ~~1269~~ ~~1270~~ ~~1271~~ ~~1272~~ ~~1273~~ ~~1274~~ ~~1275~~ ~~1276~~ ~~1277~~ ~~1278~~ ~~1279~~ ~~1280~~ ~~1281~~ ~~1282~~ ~~1283~~ ~~1284~~ ~~1285~~ ~~1286~~ ~~1287~~ ~~1288~~ ~~1289~~ ~~1290~~ ~~1291~~ ~~1292~~ ~~1293~~ <

TCP fast retransmit

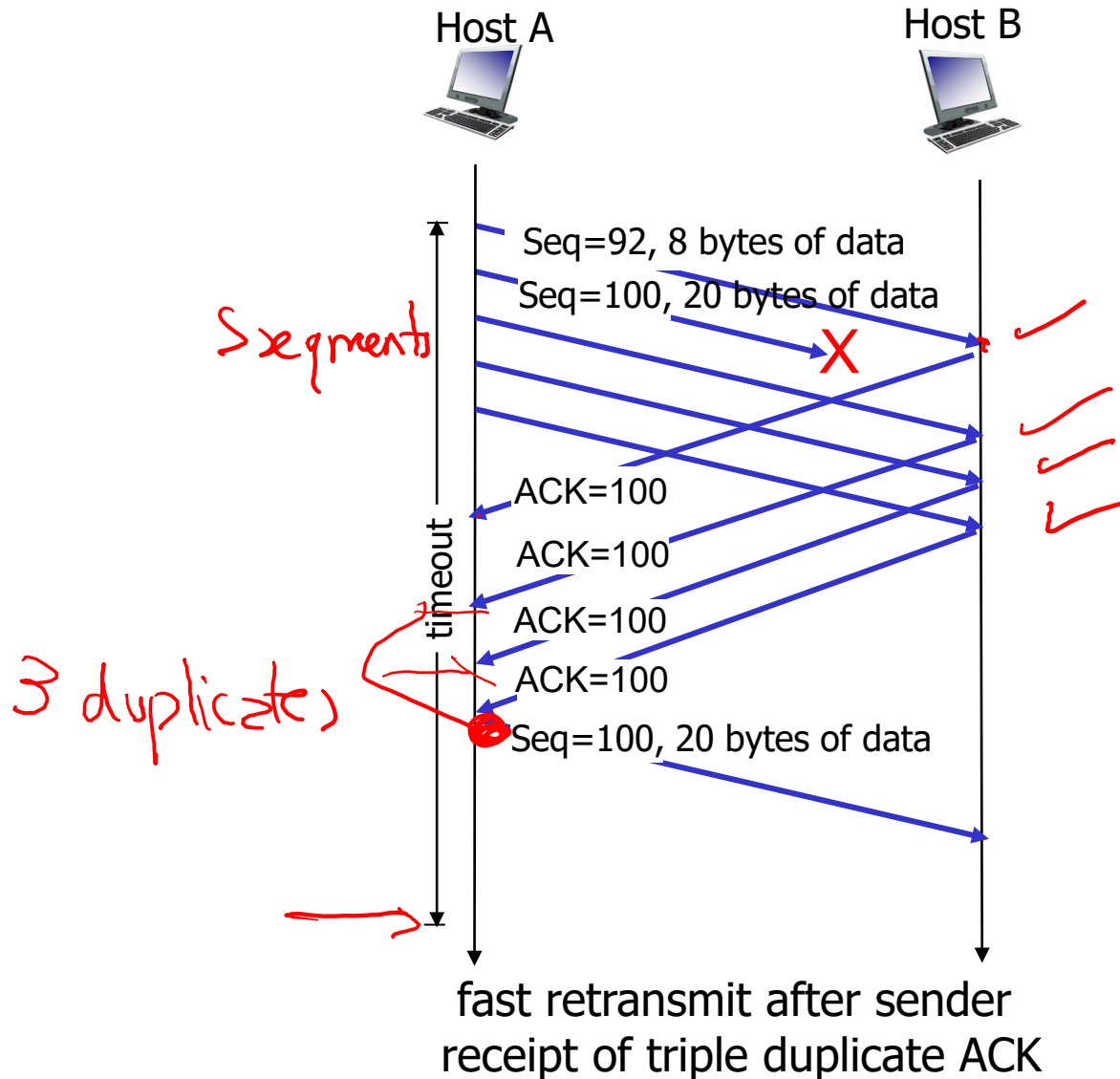
- time-out period often relatively long:
 - long delay before resending lost packet
- detect lost segments via duplicate ACKs.
 - sender often sends many segments back-to-back
 - if segment is lost, there will likely be many duplicate ACKs.

TCP fast retransmit

if sender receives 3 ACKs for same data (“triple duplicate ACKs”), resend unacked segment with smallest seq #

- likely that unacked segment lost, so don't wait for timeout

TCP fast retransmit

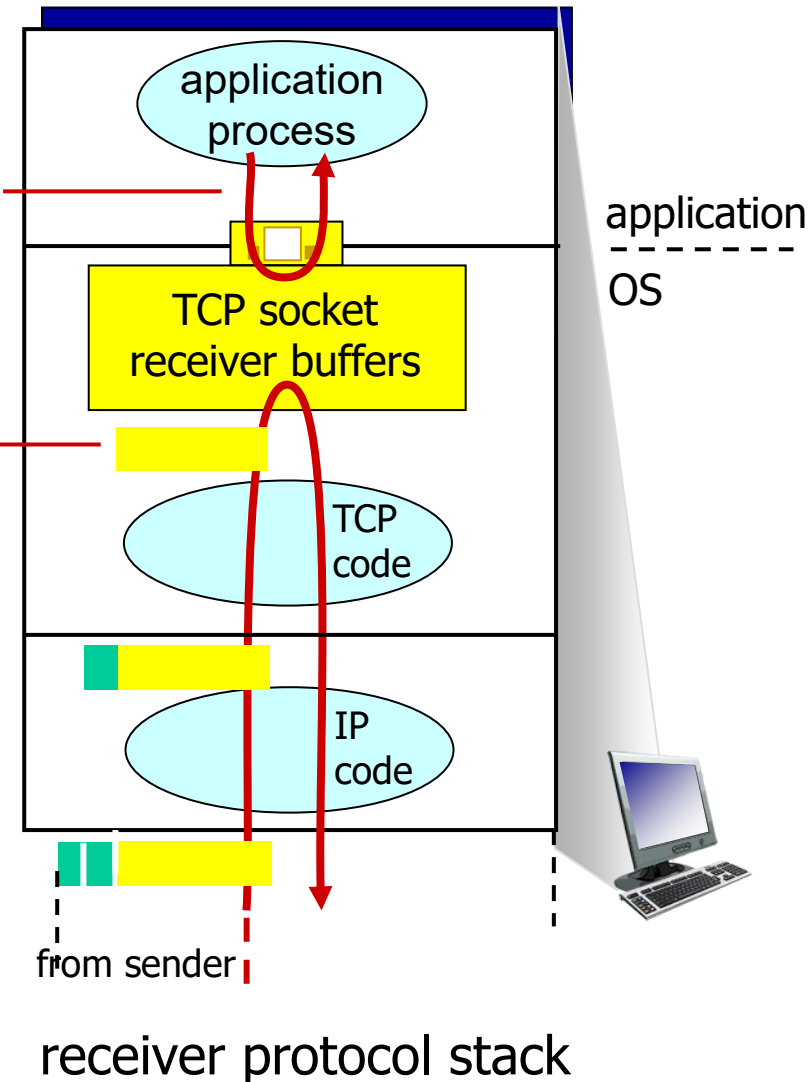


TCP flow control

application may
remove data from
TCP socket buffers

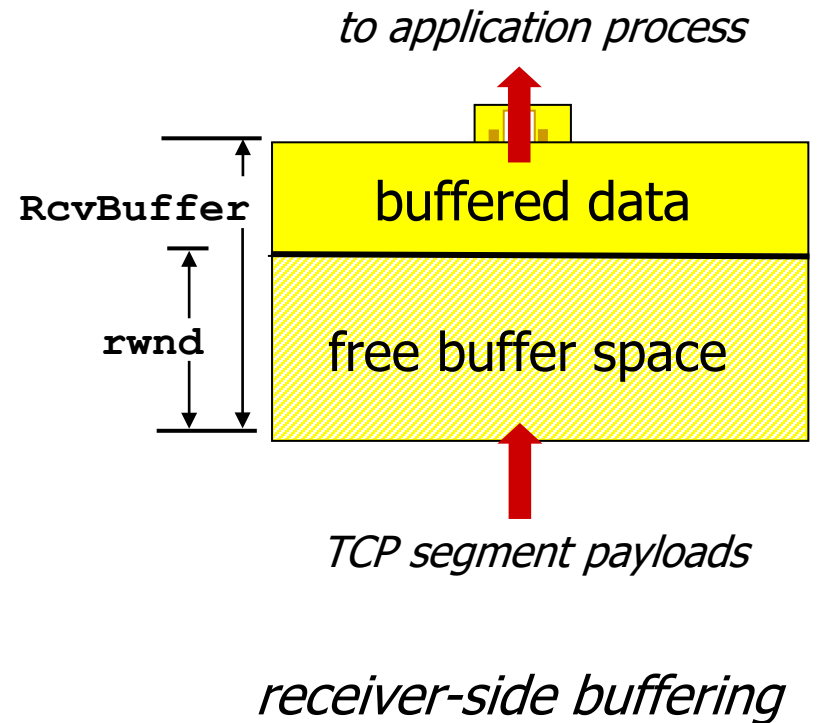
... slower than TCP
receiver is delivering
(sender is sending)

flow control
receiver controls sender, so
sender won't overflow
receiver's buffer by transmitting
too much, too fast



TCP flow control

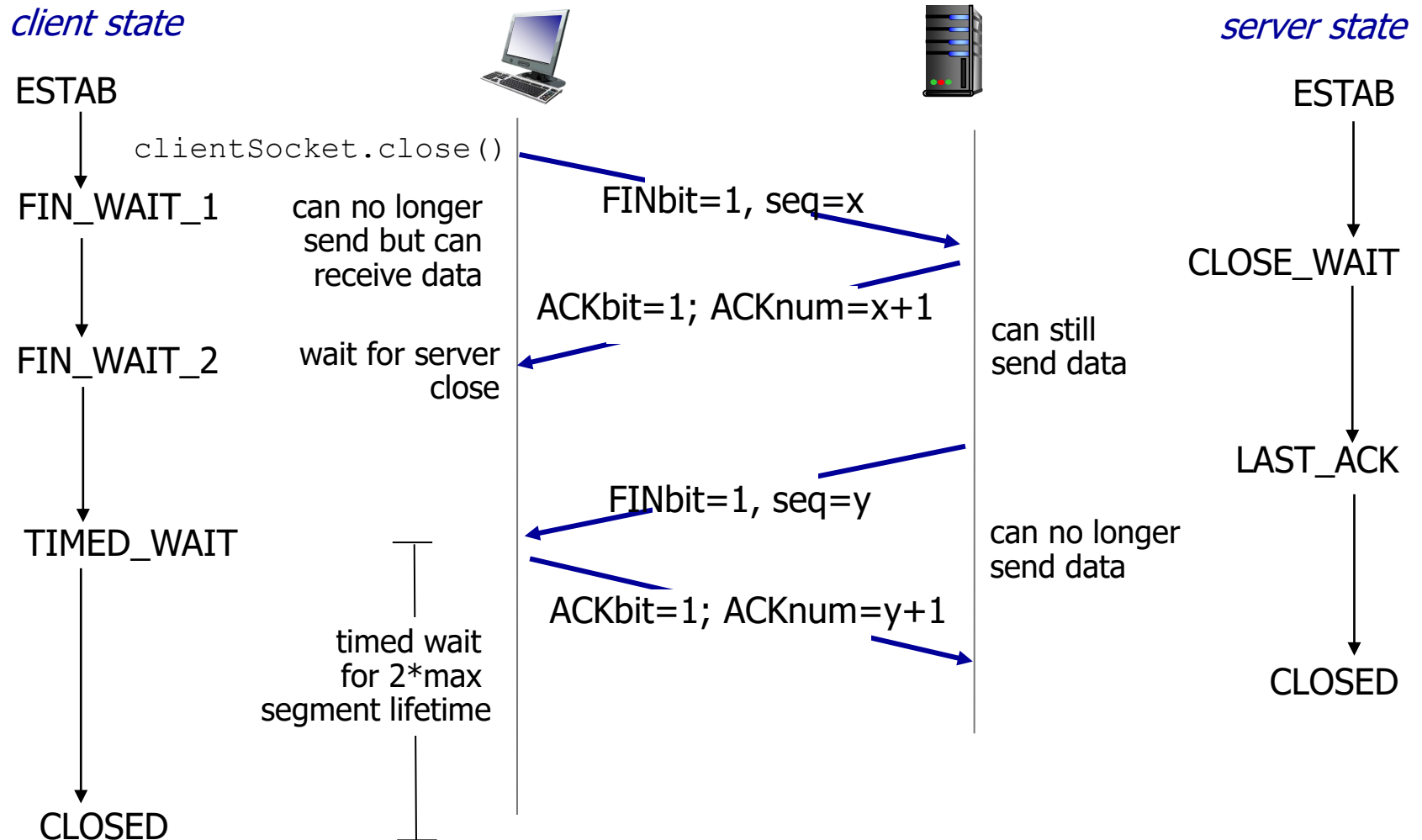
- receiver “advertises” free buffer space by including **rwnd** value in TCP header of receiver-to-sender segments
 - **RcvBuffer** size set via socket options (typical default is 4096 bytes)
 - many operating systems autoadjust **RcvBuffer**
- sender limits amount of unacked (“in-flight”) data to receiver’s **rwnd** value
- guarantees receive buffer will not overflow



TCP: closing a connection

- client, server each close their side of connection
 - send TCP segment with FIN bit = 1
- respond to received FIN with ACK
 - on receiving FIN, ACK can be combined with own FIN
- simultaneous FIN exchanges can be handled

TCP: closing a connection



Transport Layer

- ✓ Transport-Layer services
- ✓ Multiplexing and demultiplexing
- ✓ UDP
- ✓ Reliable data transfer
- ✓ TCP
- ✓ Principles of congestion control
- ✓ TCP Congestion control

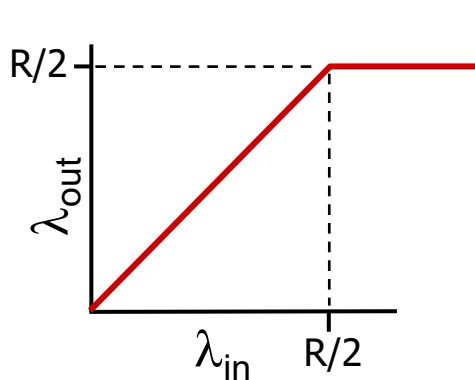
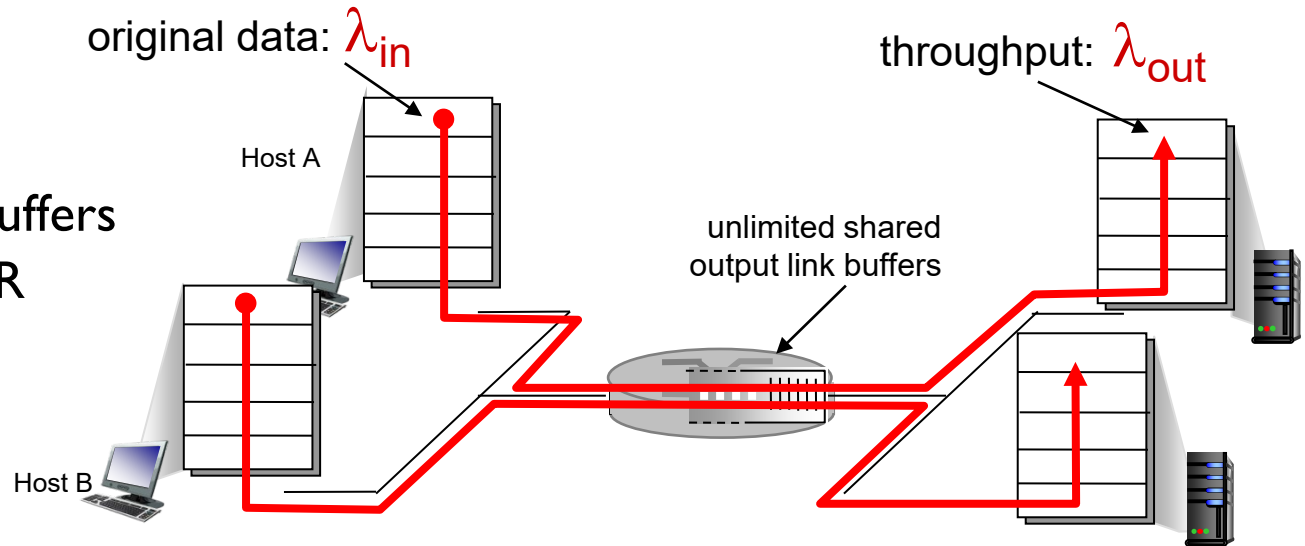
Principles of congestion control

congestion:

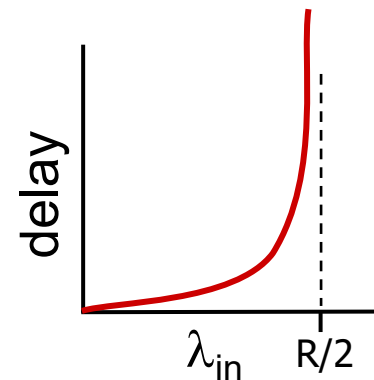
- informally: “too many sources sending too much data too fast for *network* to handle”
- different from flow control!
- manifestations:
 - lost packets (buffer overflow at routers)
 - long delays (queueing in router buffers)
- a top-10 problem!

Causes/costs of congestion: scenario I

- two senders, two receivers
- one router, infinite buffers
- output link capacity: R
- no retransmission



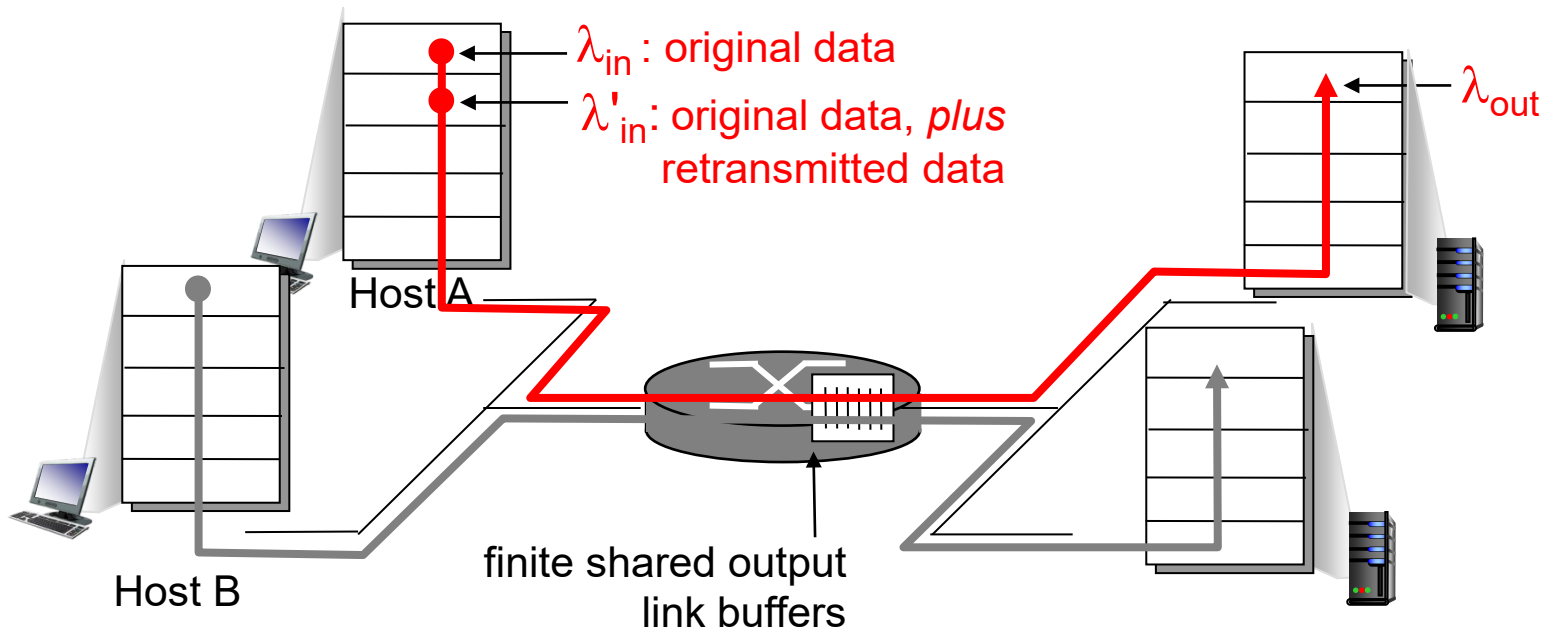
- maximum per-connection throughput: $R/2$



- ❖ large delays as arrival rate, λ_{in} , approaches capacity

Causes/costs of congestion: scenario 2

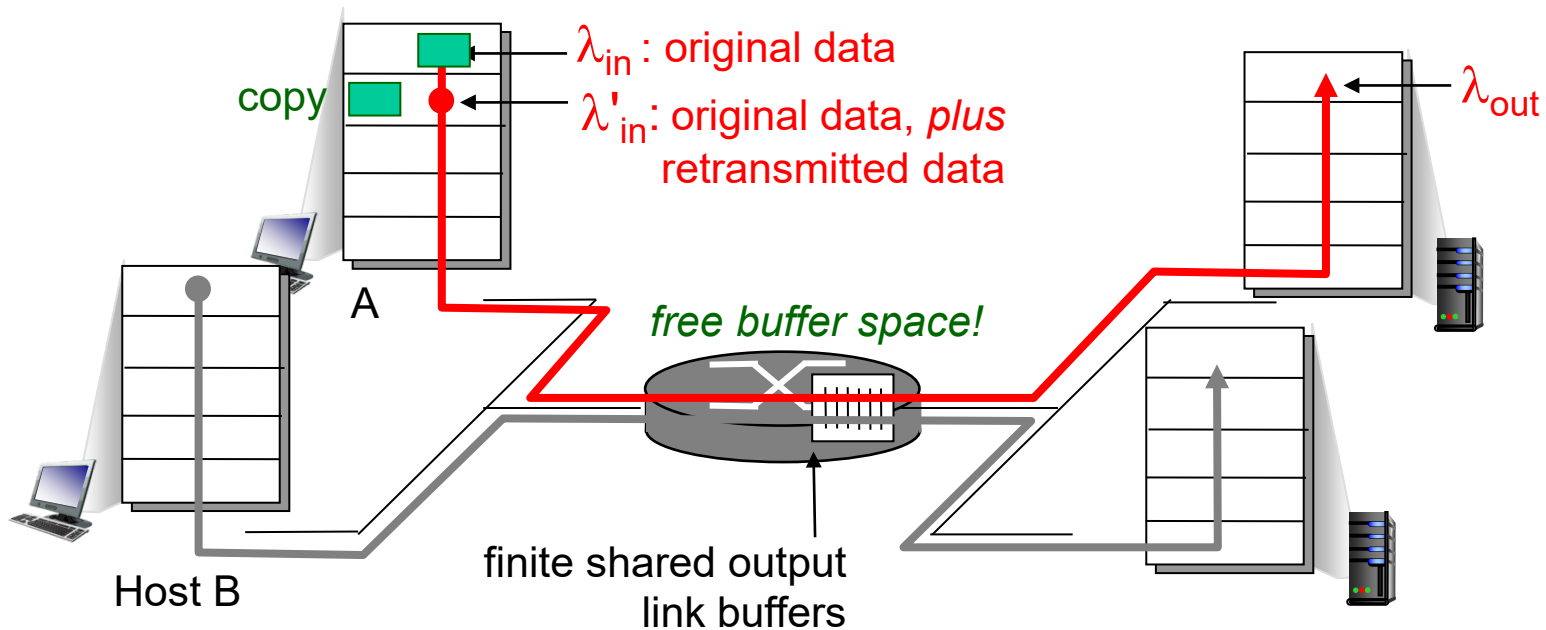
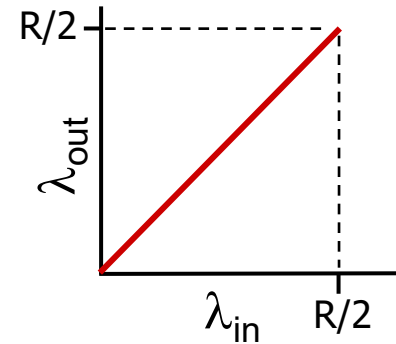
- one router, *finite* buffers
- sender retransmission of timed-out packet
 - application-layer input = application-layer output: $\lambda_{in} = \lambda_{out}$
 - transport-layer input includes *retransmissions* : $\lambda'_{in} \geq \lambda_{in}$



Causes/costs of congestion: scenario 2

idealization: perfect knowledge

- sender sends only when router buffers available

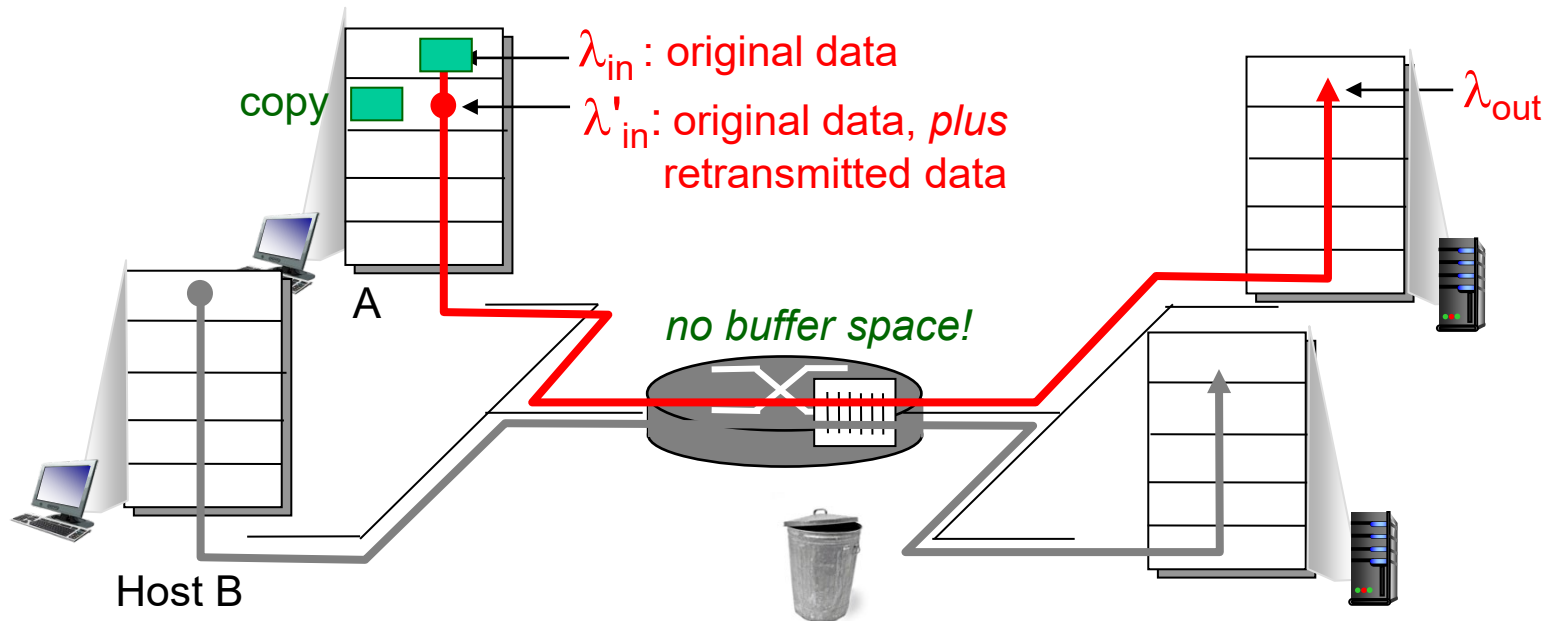


Causes/costs of congestion: scenario 2

Idealization: known loss

packets can be lost,
dropped at router due
to full buffers

- sender only resends if
packet *known* to be lost

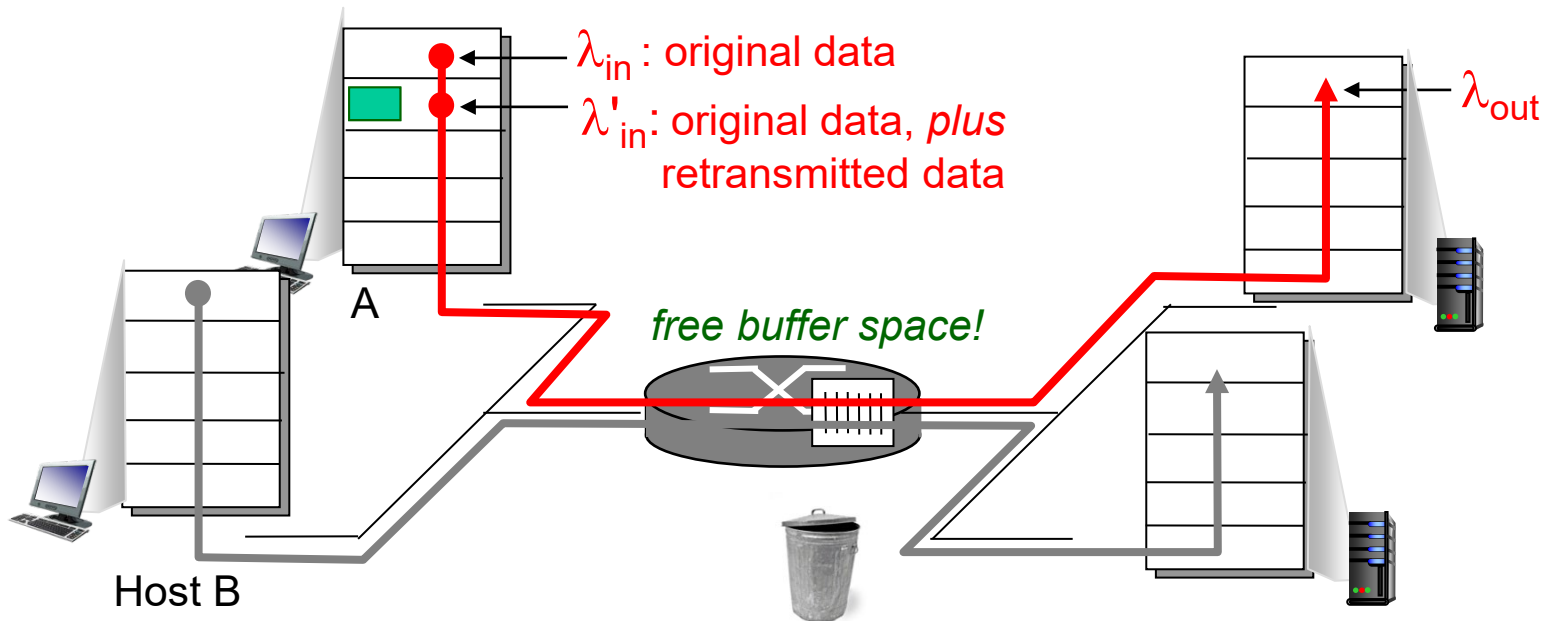
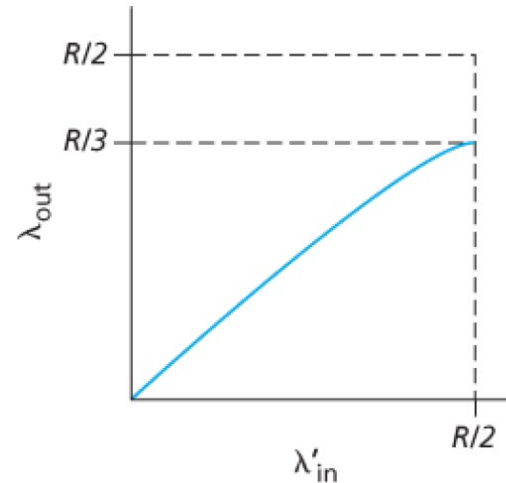


Causes/costs of congestion: scenario 2

Idealization: known loss

packets can be lost,
dropped at router due
to full buffers

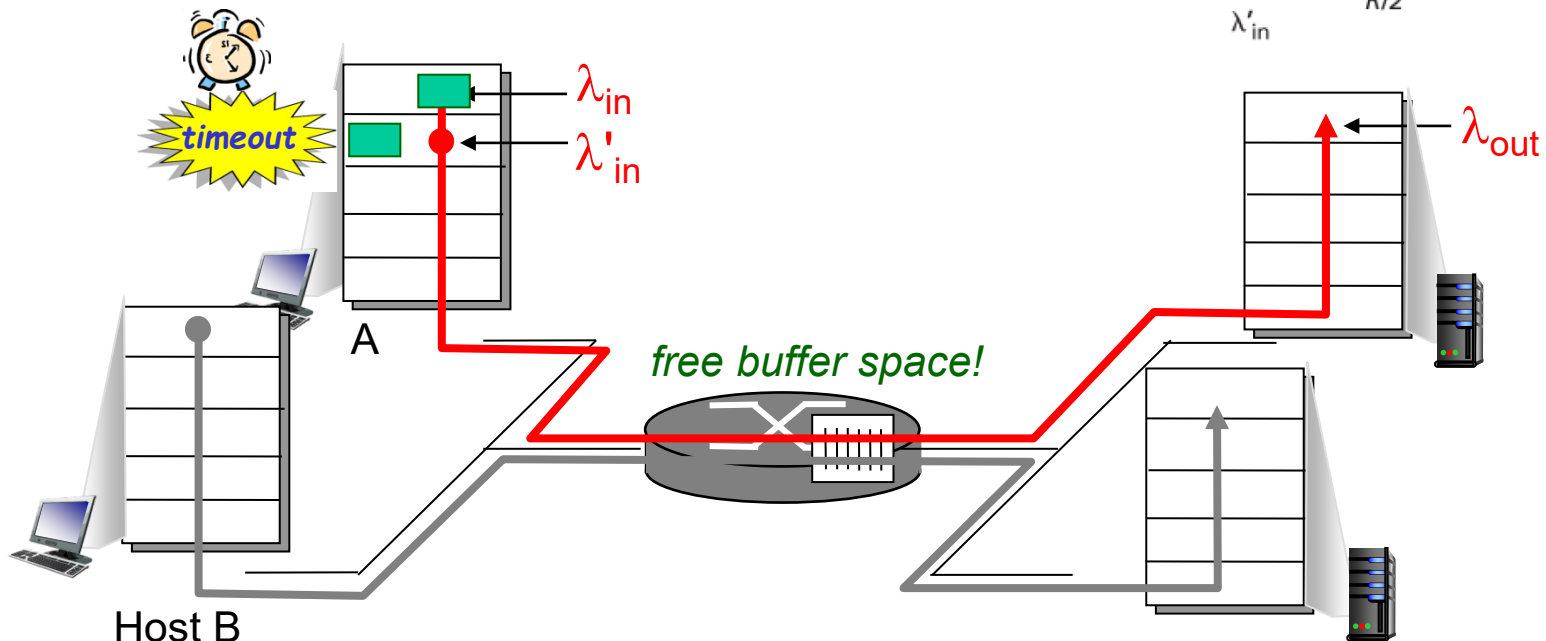
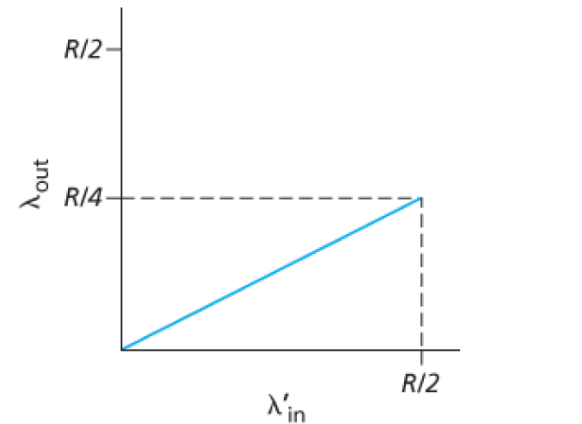
- sender only resends if
packet *known* to be lost



Causes/costs of congestion: scenario 2

Realistic: *duplicates*

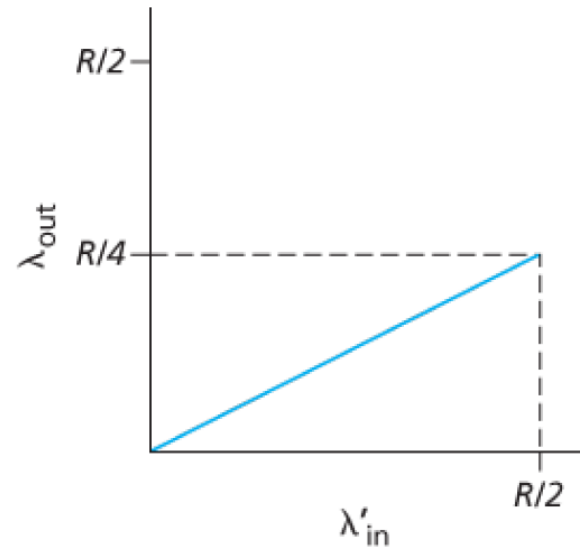
- packets can be lost, dropped at router due to full buffers
- sender times out prematurely, sending *two* copies, both of which are delivered



Causes/costs of congestion: scenario 2

Realistic: duplicates

- packets can be lost, dropped at router due to full buffers
- sender times out prematurely, sending *two* copies, both of which are delivered



“costs” of congestion:

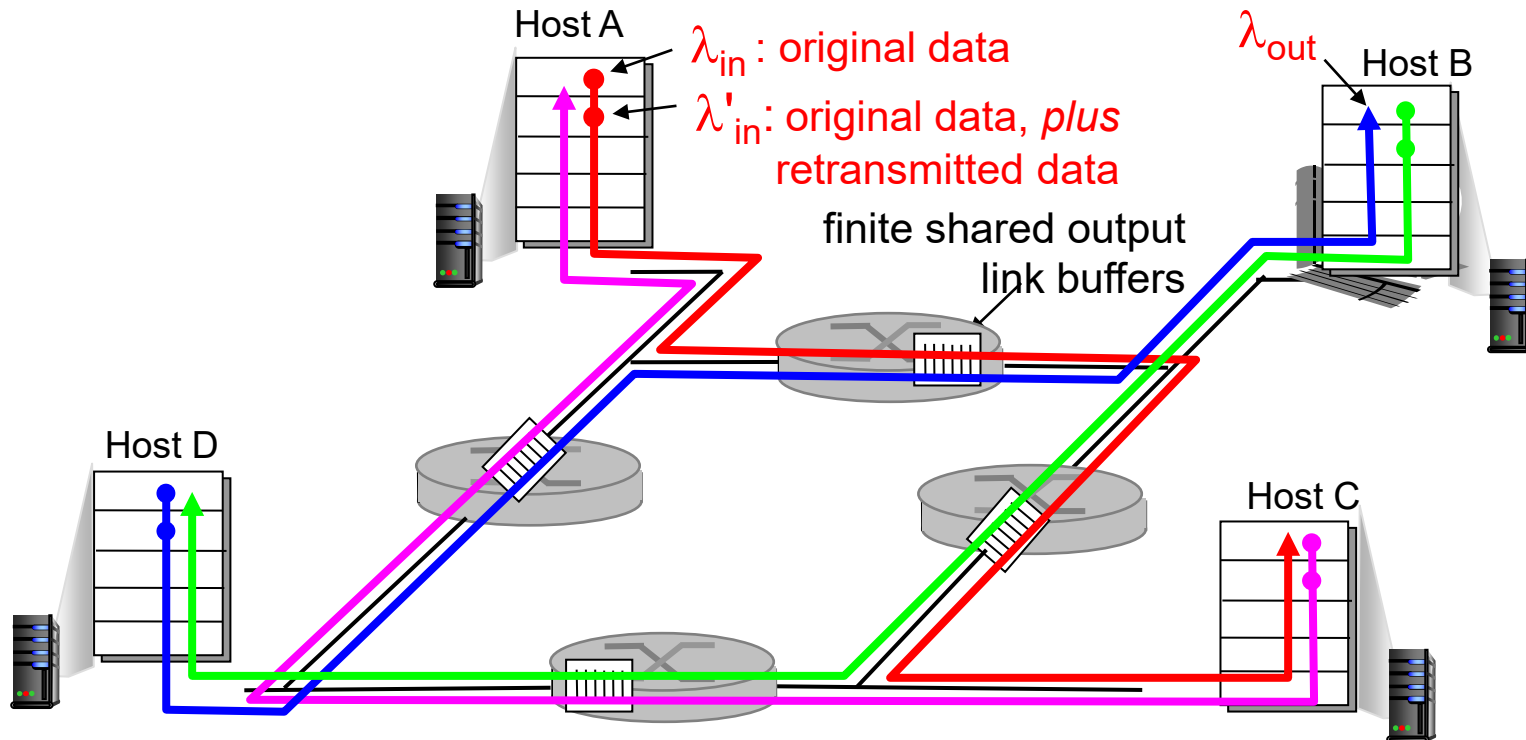
- more work (retrans) for given “goodput”
- unneeded retransmissions: link carries multiple copies of pkt
 - decreasing goodput

Causes/costs of congestion: scenario 3

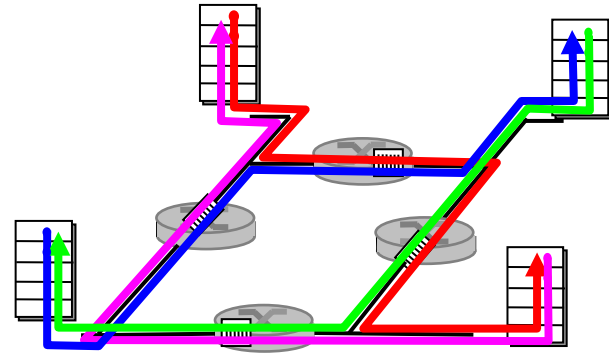
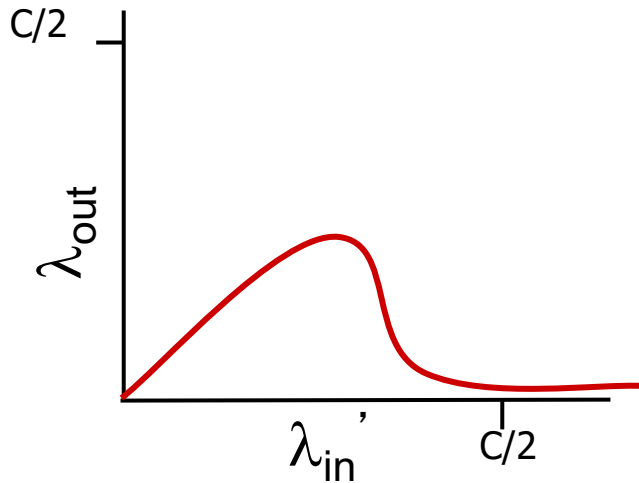
- four senders
- multihop paths
- timeout/retransmit

Q: what happens as λ_{in} and λ'_{in} increase ?

A: as red λ'_{in} increases, all arriving blue pkts at upper queue are dropped, blue throughput $\rightarrow 0$



Causes/costs of congestion: scenario 3



another “cost” of congestion:

- when packet dropped, any “upstream transmission capacity used for that packet was wasted!

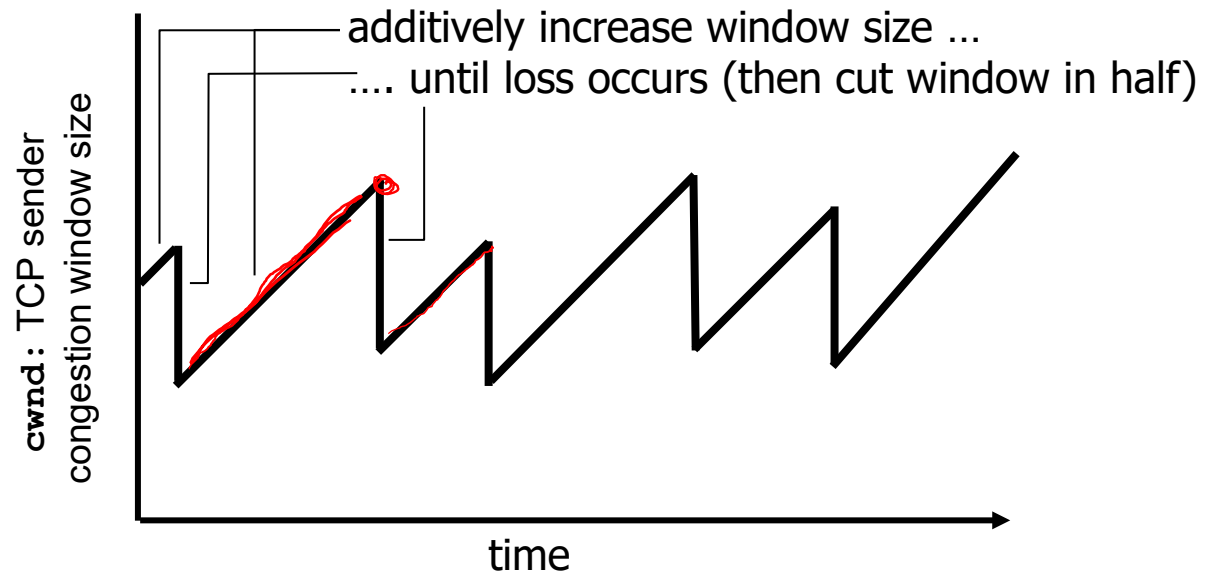
Transport Layer

- ✓ Transport-Layer services
- ✓ Multiplexing and demultiplexing
- ✓ UDP
- ✓ Reliable data transfer
- ✓ TCP
- ✓ Principles of congestion control
- ✓ TCP Congestion control

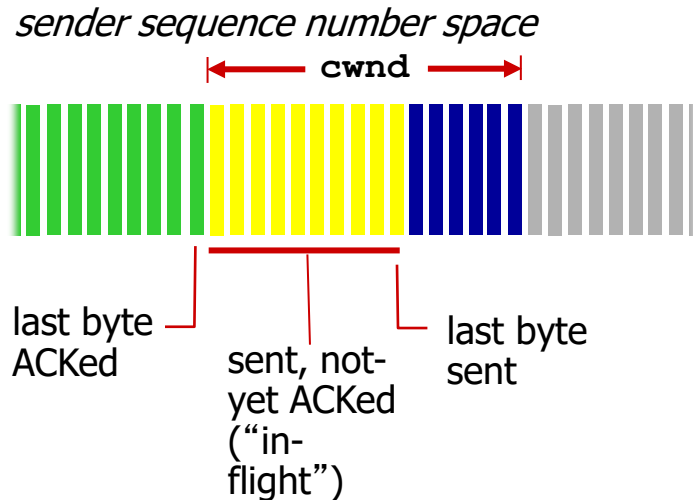
TCP congestion control: additive increase multiplicative decrease

- *approach*: sender increases transmission rate (window size), probing for usable bandwidth, until loss occurs
 - *additive increase*: increase **cwnd** by 1 MSS every RTT until loss detected
 - *multiplicative decrease*: cut **cwnd** in half after loss

AIMD saw tooth
behavior: probing
for bandwidth



TCP Congestion Control: details



- sender limits transmission:

$$\text{LastByteSent} - \text{LastByteAcked} \leq \text{cwnd}$$

- **cwnd** is dynamic, function of perceived network congestion

TCP sending rate:

- roughly: send cwnd bytes, wait RTT for ACKS, then send more bytes

$$\text{rate} \approx \frac{\text{cwnd}}{\text{RTT}} \text{ bytes/sec}$$

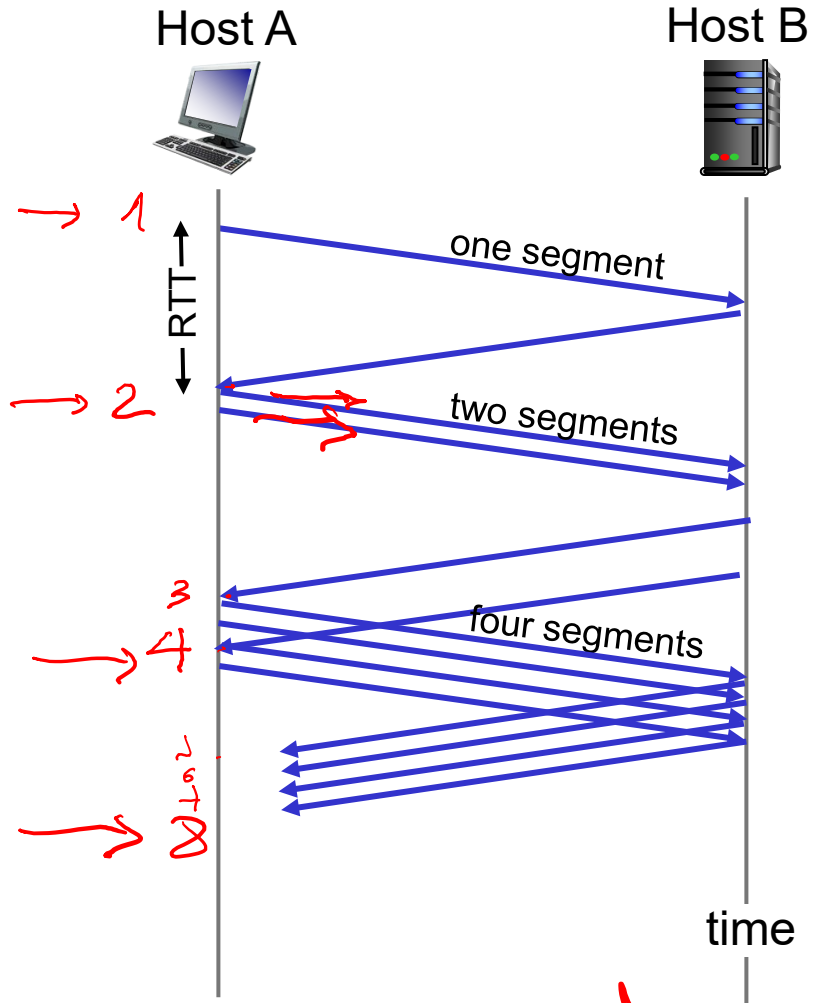
min(rcwnd, cwnd)
✓ Slow start
✓ Congestion avoidance
✓ Fast recovery

TCP Slow Start

- when connection begins, increase rate exponentially until first loss event:

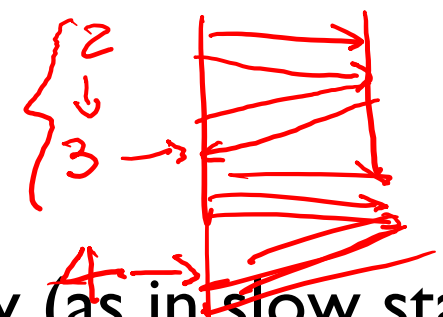
- initially **cwnd = 1 MSS**
- double **cwnd** every RTT
- done by incrementing **cwnd** for every ACK received

- summary: initial rate is slow but ramps up exponentially fast



after threshold. → Congestion avoidance

TCP: detecting, reacting to loss

- loss indicated by timeout: MSS 
 - cwnd set to 1 MSS;
 - window then grows exponentially (as in slow start) to threshold, then grows linearly → Cong. Avoidance
- loss indicated by 3 duplicate ACKs: TCP RENO
 - dup ACKs indicate network capable of delivering some segments → fast recovery
 - cwnd is cut in half window then grows linearly
- TCP Tahoe always sets cwnd to 1 (timeout or 3 duplicate acks) → slow start

slow start

cong. avoidance

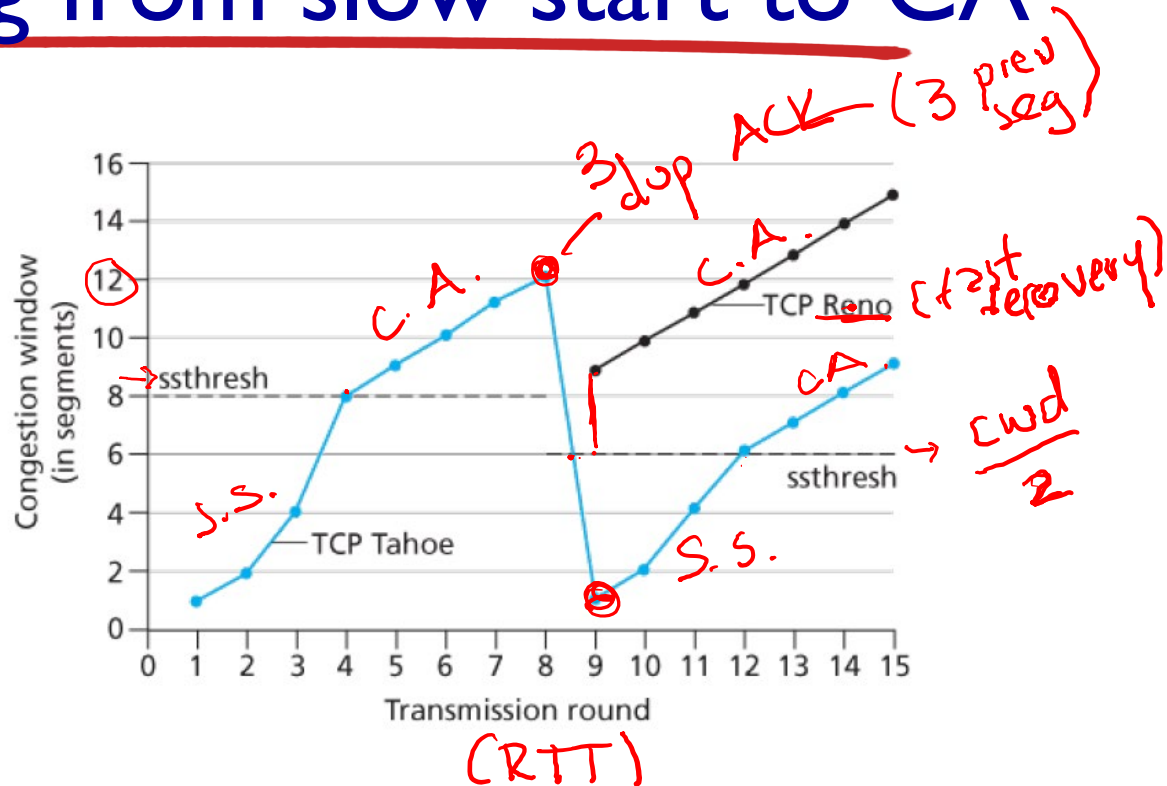
TCP: switching from slow start to CA

Q: when should the exponential increase switch to linear?

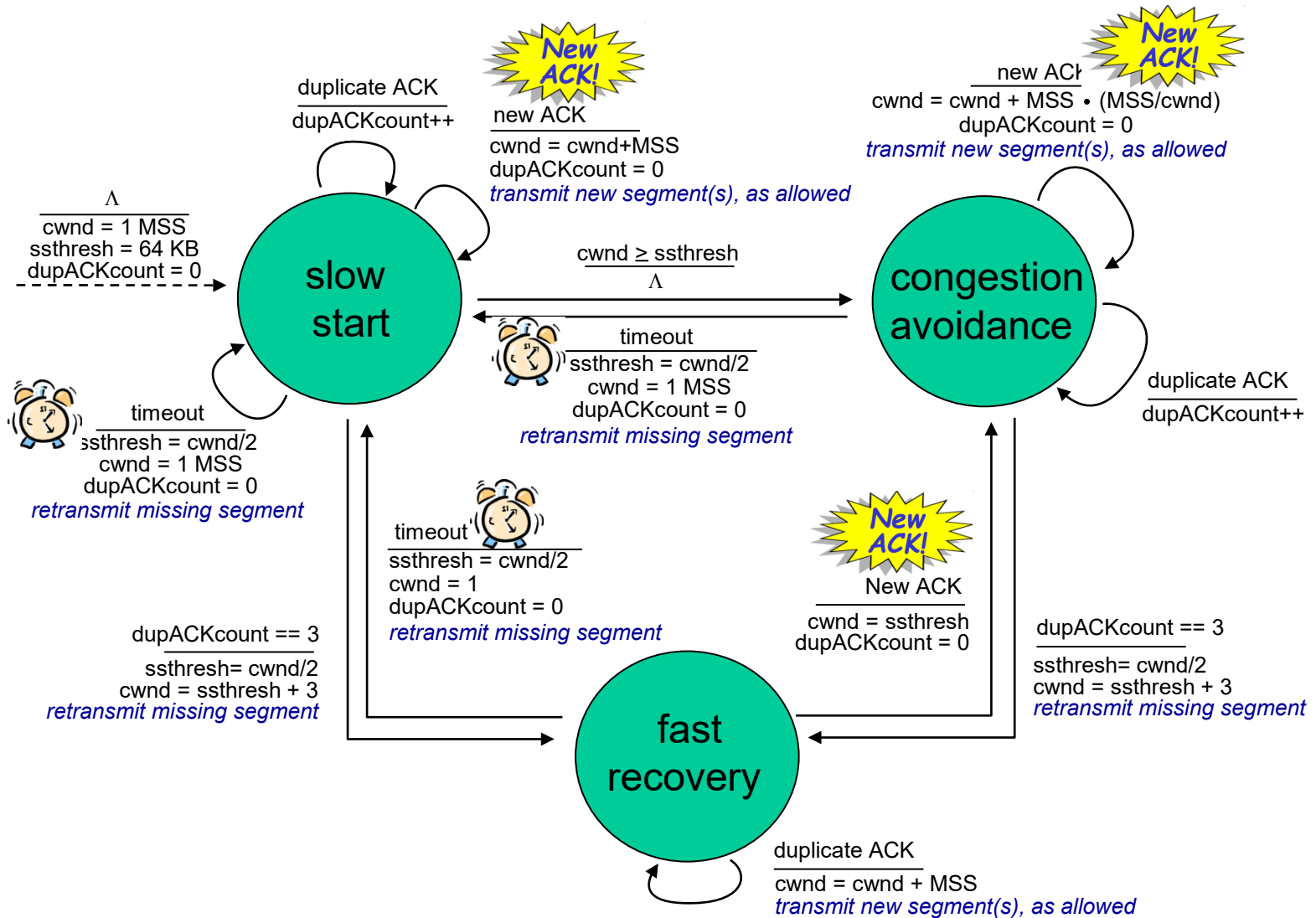
A: when **cwnd** gets to 1/2 of its value before timeout.

Implementation:

- variable **ssthresh**
- on loss event, **ssthresh** is set to 1/2 of **cwnd** just before loss event



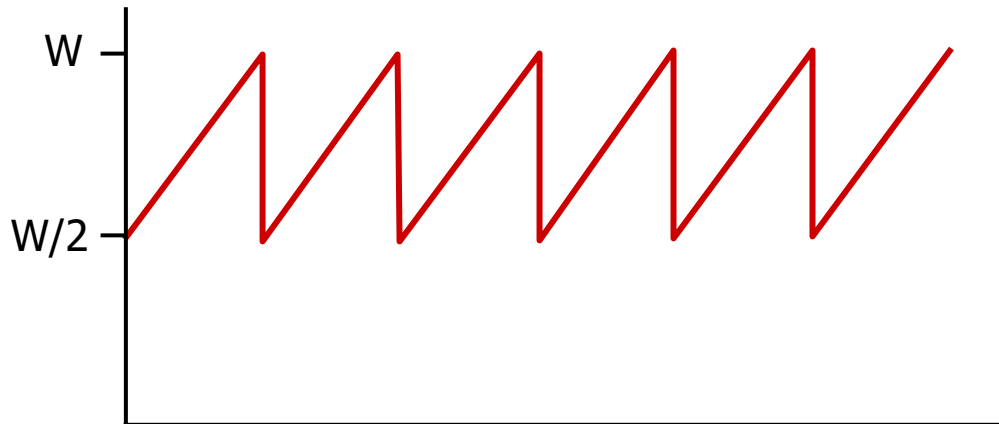
Summary: TCP Congestion Control



TCP throughput

- avg. TCP thruput as function of window size, RTT?
 - ignore slow start, assume always data to send
- **W: window size** (measured in bytes) where loss occurs
 - avg. window size (# in-flight bytes) is $\frac{3}{4} W$
 - avg. thruput is $\frac{3}{4}W$ per RTT

$$\text{avg TCP thruput} = \frac{3}{4} \frac{W}{RTT} \text{ bytes/sec}$$



$$\frac{\frac{W}{RTT} + \frac{W}{2RTT}}{2}$$

TCP Futures: TCP over “long, fat pipes”

- example: 1500 byte segments, 100ms RTT, want 10 Gbps throughput
- requires $W = 83,333$ in-flight segments
- throughput in terms of segment loss probability, L [Mathis 1997]:

$$\text{TCP throughput} = \frac{1.22 \cdot \text{MSS}}{\text{RTT} \sqrt{L}}$$

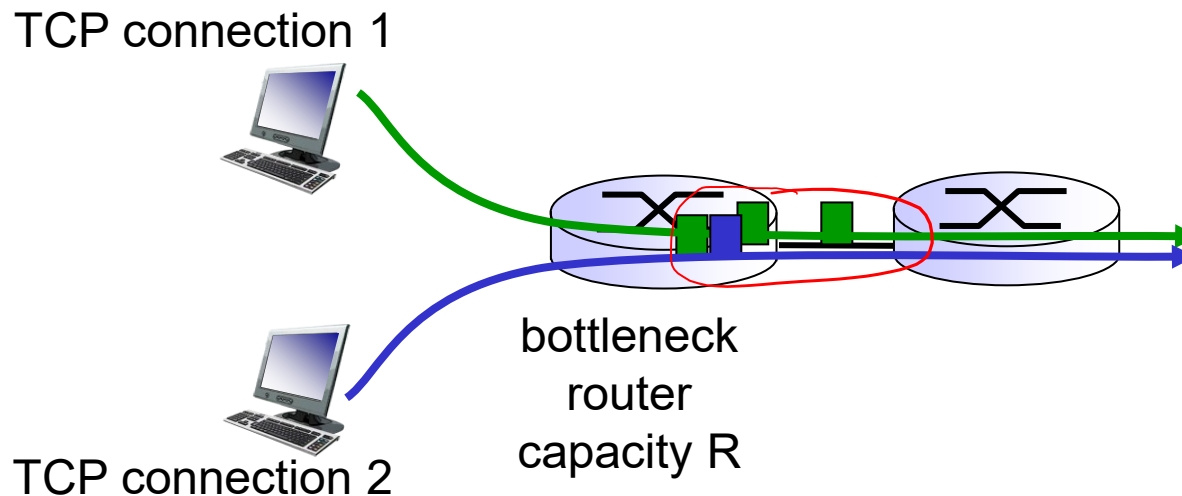
packet loss prop.

→ to achieve 10 Gbps throughput, need a loss rate of $L = 2 \cdot 10^{-10}$ — *a very small loss rate!*

- new versions of TCP for high-speed

TCP Fairness

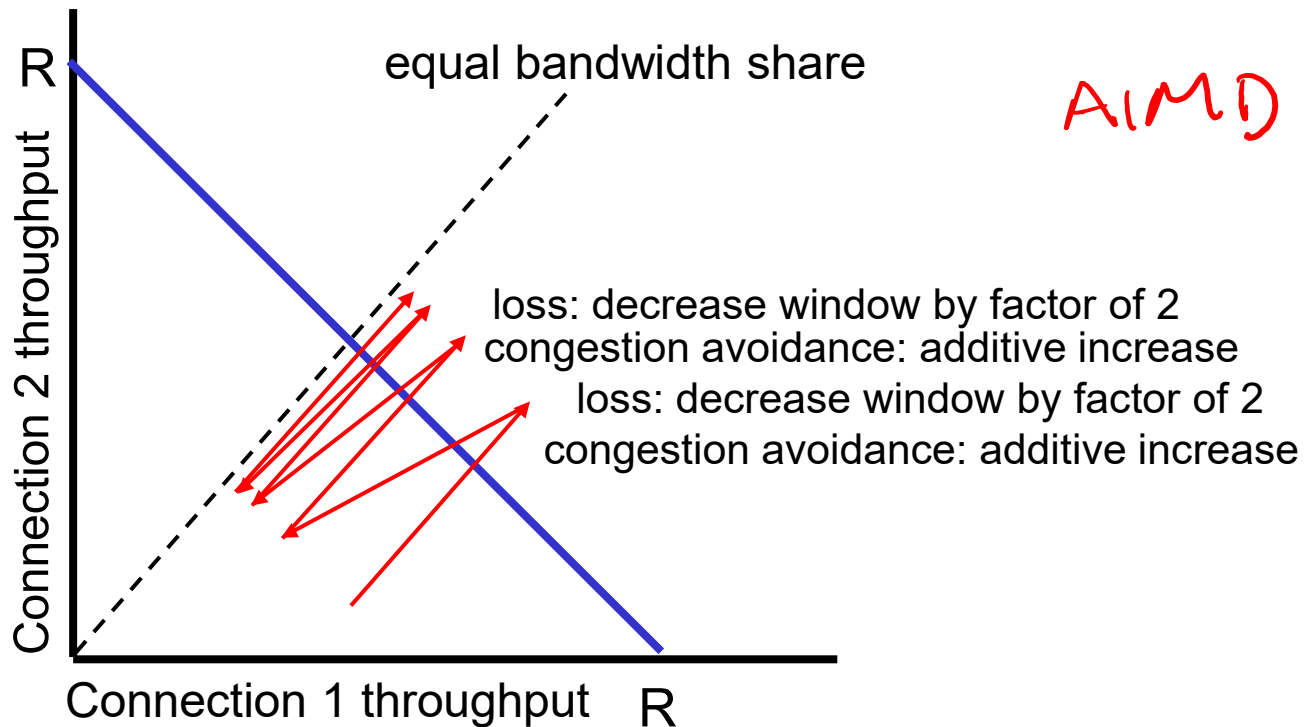
fairness goal: if K TCP sessions share same bottleneck link of bandwidth R , each should have average rate of R/K



Why is TCP fair?

two competing sessions:

- additive increase gives slope of 1, as throughput increases
- multiplicative decrease decreases throughput proportionally



Fairness (more)

Fairness and UDP

- multimedia apps often do not use TCP
 - do not want rate throttled by congestion control
- instead use UDP:
 - send audio/video at constant rate, tolerate packet loss

Fairness, parallel TCP connections

- application can open multiple parallel connections between two hosts
- web browsers do this
- e.g., link of rate R with 9 existing connections:
 - new app asks for 1 TCP, gets rate $R/10$
 - new app asks for 11 TCPs, gets $R/2$

Chapter 3: summary

- principles behind transport layer services:
 - multiplexing, demultiplexing
 - reliable data transfer
 - flow control
 - congestion control
- instantiation, implementation in the Internet
 - UDP
 - TCP

next:

- leaving the network “edge” (application, transport layers)
- into the network “core”
- two network layer chapters:
 - data plane
 - control plane

References

Figures and slides are taken/adapted from:

- Jim Kurose, Keith Ross, "Computer Networking: A Top-Down Approach", 7th ed. Addison-Wesley, 2012. All material copyright 1996-2016 J.F Kurose and K.W. Ross, All Rights Reserved