TCP over Wireless Networks

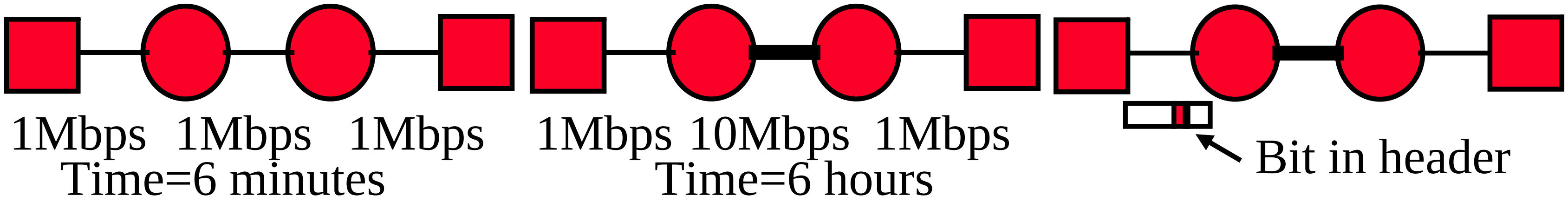
Raj Jain
Professor of Computer Science and Engineering
Washington University in Saint Louis
Saint Louis, MO 63130

Audio/Video recordings of this lecture are available at:

http://www.cse.wustl.edu/~jain/cse574-10/

- TCP Congestion Mechanisms
- Our Initial Research on TCP Congestion
- TCP Over Wireless: Issues and Solutions
- TCP over Satellite
- Our research on TCP over Satellite and Wireless

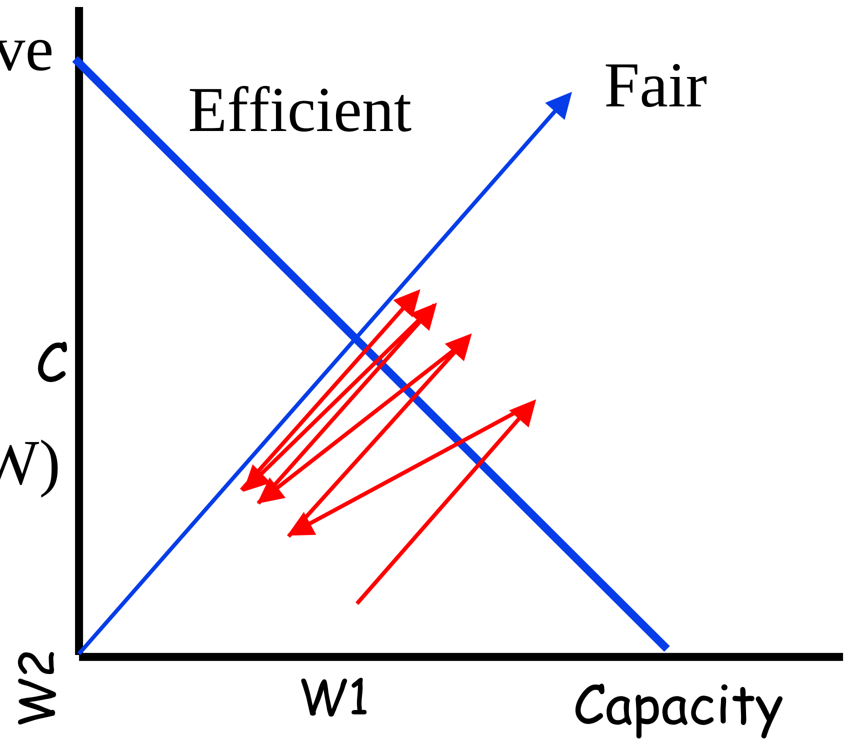
Research on Congestion Control



- Early 1980s Digital Equipment Corporation (DEC) introduced Ethernet products
- Noticed that throughput goes down with a higher-speed link in middle (because no congestion mechanisms in TCP)
- Results:
  1. Timeout ⇒ Congestion
     ⇒ Reduce the TCP window to one on a timeout [Jain 1986]
  2. Routers should set a bit when congested (DECbit).
     [Jain, Ramakrishnan, Chiu 1988]
  3. Introduced the term "Congestion Avoidance"
  4. Additive increase and multiplicative decrease (AIMD principle)
     [Chiu and Jain 1989]
- There were presented to IETF in 1986.
  ⇒ Slow-start based on Timeout and AIMD [Van Jacobson 1988]

AIMD Principle

- Additive Increase, Multiplicative Decrease
- W1+W2 = Capacity
  ⇒ Efficiency,
  W1=W2 ⇒ Fairness
- (W1,W2) to (W1+DW,W2+DW)
  ⇒ Linear increase (45° line)
- (W1,W2) to (kW1,kW2)
  ⇒ Multiplicative decrease
  (line through origin)



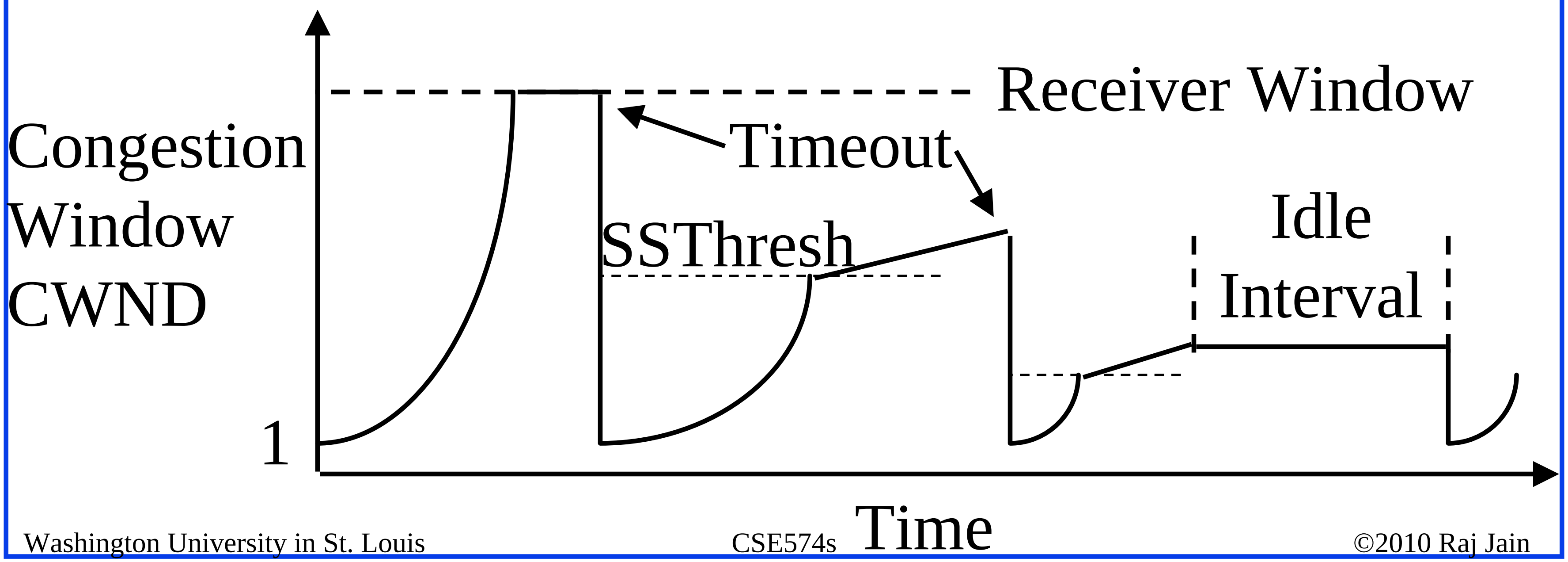
Ref: D. Chiu and Raj Jain, "**Analysis of the Increase/Decrease Algorithms for Congestion Avoidance in Computer Networks**," Journal of Computer Networks and ISDN, Vol. 17, No. 1, June 1989, pp. 1-14, http://www.cse.wustl.edu/~jain/papers/cong_av.htm

Slow Start Congestion Control

- Window = Flow Control Avoids receiver overrun
- Need congestion control to avoid network overrun
- The sender maintains two windows:
  Credits from the receiver
  Congestion window from the network
  Congestion window is always less than the receiver window
- Starts with a congestion window (CWND) of 1 segment (one max segment size)
  ⇒ Do not disturb existing connections too much.
- Increase CWND by 1 MSS every time an ack is received

Slow Start (Cont)

- If segments lost, remember slow start threshold (SSThresh) to CWND/2
  Set CWND to 1 MSS
  Increment by 1 per ack until SSthresh
  Increment by 1 MSS/CWND per ack afterwards



Slow Start (Cont)

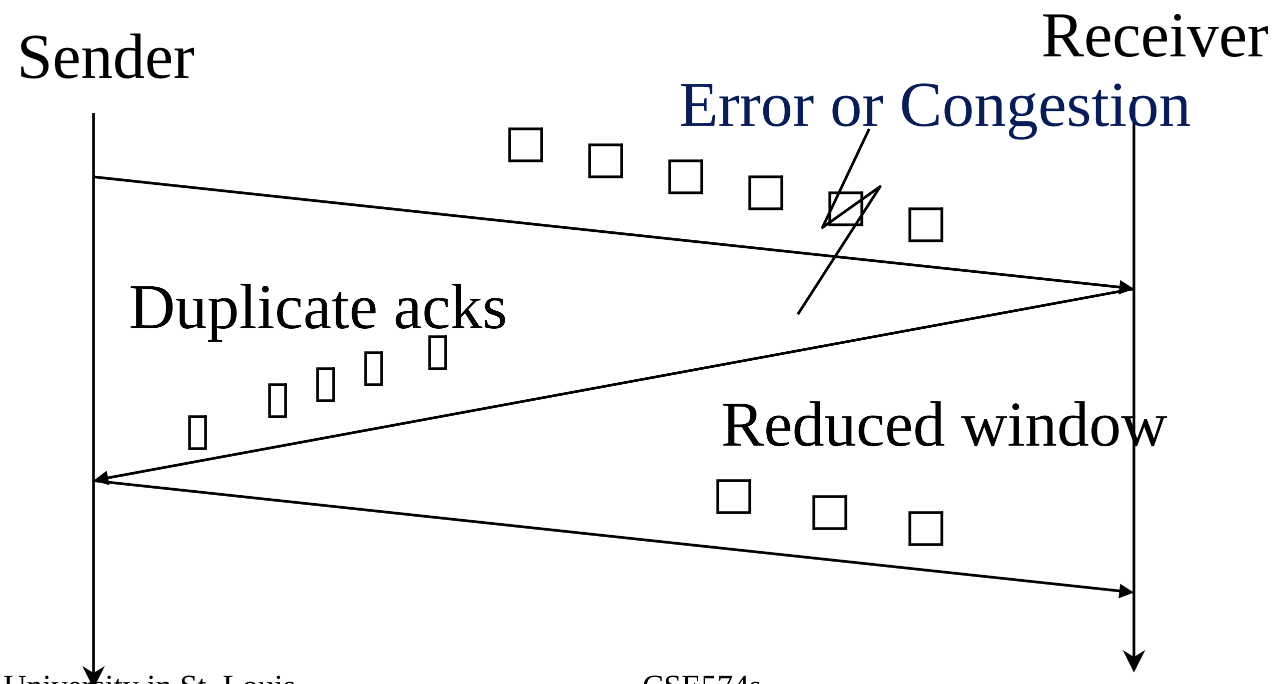
- At the beginning, SSThresh = Receiver window
- After a long idle period (exceeding one round-trip time), reset the congestion window to one.
- Exponential growth phase is also known as "Slow start" phase
- The linear growth phase is known as "congestion avoidance phase"

Fast Recovery

- Optional – implemented in TCP Reno
  (Earlier version was TCP Tahoe)
- Duplicate Ack indicates a lost/out-of-order segment
- On receiving 3 duplicate acks:
  - Enter Fast Recovery mode
    - Retransmit missing segment
    - Set SSTHRESH=CWND/2
    - Set CWND=SSTHRESH+3 MSS
    - Every subsequent duplicate ack: CWND=CWND+1MSS

Problems of Current TCP

- TCP cannot distinguish wireless errors from congestion.
- Frequent errors ⇒ Frequent window reductions
  ⇒ Low throughput
- On CDMA, Overload ⇒ Errors. Otherwise no relationship.

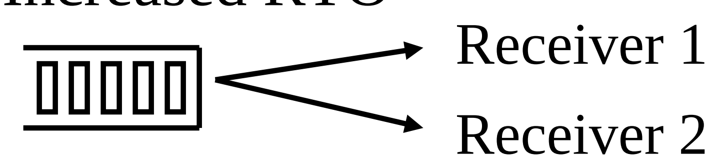


TCP Over Wireless

1. Link Layer Mechanisms
2. Split TCP Solutions
3. TCP Aware Link Layer Protocols
4. Explicit Notification Schemes
5. TCP Over Satellite
6. Our Results for Satellite and Wireless Networks

1. Link Layer Mechanisms

- Forward Error Correction (FEC):
  - Reduces loss due to errors.
  - Reduced link throughput even if no errors.
- Automatic Repeat Request (ARQ):
  - Link layer retransmission and acknowledgement
  - No reduction in throughput if no errors
  - Reduced throughput and increased delay at link layer
  - May cause congestion
  - May increase variance of RTT $\Rightarrow$ Increased RTO
  - May cause head-of-line blocking

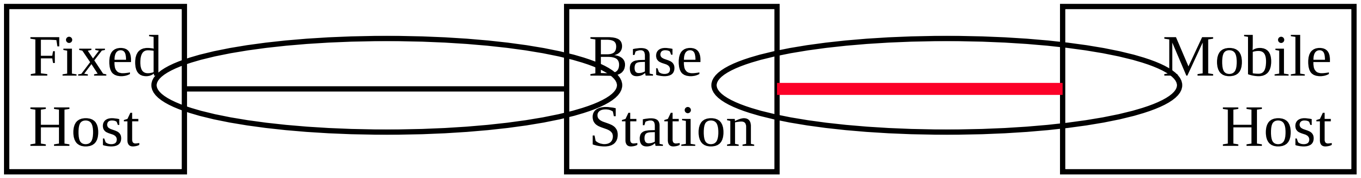


- Adaptive Link layer strategies:
  - Dynamically vary FEC code, retransmission limit, frame size

2. Split TCP Solutions

- **Indirect TCP**
- Selective Repeat Protocol (SRP)
- Mobile TCP
- Mobile-End Transport Protocol

Indirect TCP



- Two TCP connections:
  - Fixed host to Base
  - Base to Mobile
- Independent flow control on two connections
- Packets buffered in the base
- Ack at sender ≠ MH has received
  - Violates TCP's end-to-end semantics
  - BS retains hard state. BS failure ⇒ loss of data
  - On handoff, stored packets must be sent to new BS
  - Does not work if connection not bi-directional. E.g., satellites
- Reference: Bakre95, Bakre 97

3. TCP Aware Link Layer Protocols

- Snoop Protocol
- WTCP
- Delayed DupAcks Protocol
- SCPS-TP

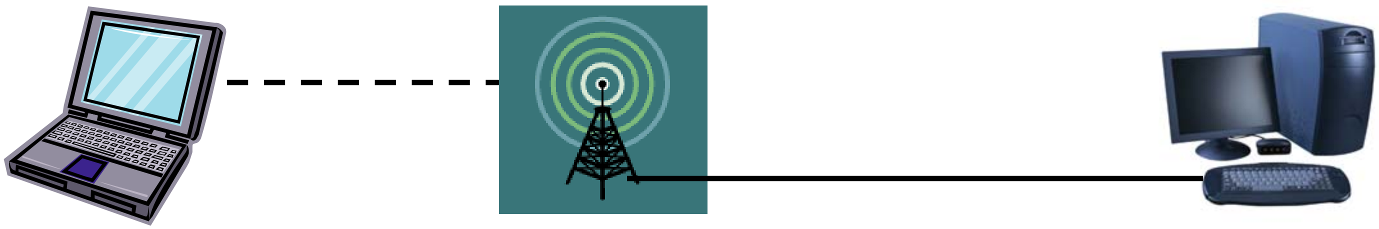
Snoop Protocol

- Split connection and link level retransmission
- Base monitors returning acks. Retransmits on duplicate acks and drops the duplicate ack
- Advantages: Only soft state at BS. Only BS modified. No changes to FH or MH.
- If wireless link delay is less than 4 packets, 3 duplicate acks will not happen and a simple link-level retransmission without dropping duplicate ack will also work.
- Disadvantages: Does not work with encrypted packets
- Does not work on asymmetric paths
- Ref: Balakrishnan95

4. Explicit Notification Schemes

- Explicit Loss Notification
- Explicit Loss Notification 2
- Explicit Bad State Notification
- Partial Ack Protocols

Explicit Loss Notification



- Works with Mobile host sources
  First link on the path is wireless
- BS keeps track of missing packets from mobile
- When DupAcks is received, BS sets "ELN" bit in the DupAcks
- When mobile receives the DupAcks with ELN bit, it does not back off. Simply retransmits.

- Reference: Balakrishnan98

5. TCP Over Satellite

- IETF TCPSAT
- Satellite Transport Protocol (STP)
- Early Acks: ACKprime

IETF TCPSAT

- Large propagation delays => Large bandwidth delay product => Large windows => Use window scaling option

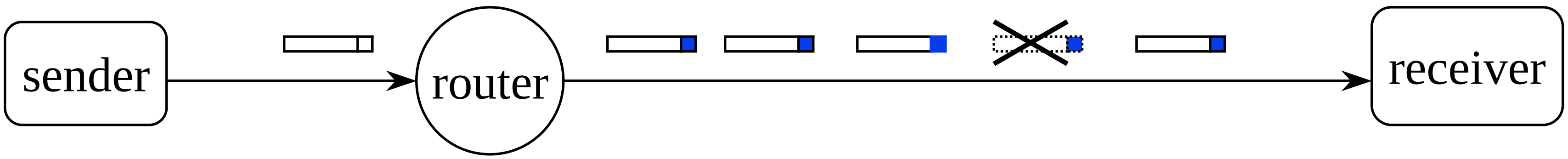
$$\text{Window} = \text{window} * 2^{\text{Scaling factor}}$$

- Use Selective acknowledgements
  => Allows multiple packets to recovered in one RTT
- Do not delay acks => Ack every packet
- Use larger initial window size (suggested 4kB)
- Byte Counting: Increase window by number of bytes acked rather than just 1 MSS per RTT
- Reduce bursts from the sender
- Ref: RFC 2488, 2760

6. Our Results for Satellite Networks

- End System Improvements:
  - Slow start
  - Fast Retransmit and Recovery
  - New Reno
  - SACK
- Intermediate System Improvements: Drop policies
- For satellite paths, end system improvements have more impact than intermediate-system based improvements
- SACK helps significantly
- Fairness depends upon the drop policies in the intermediate systems and not on end system policies

Wireless Networks: Our Solution



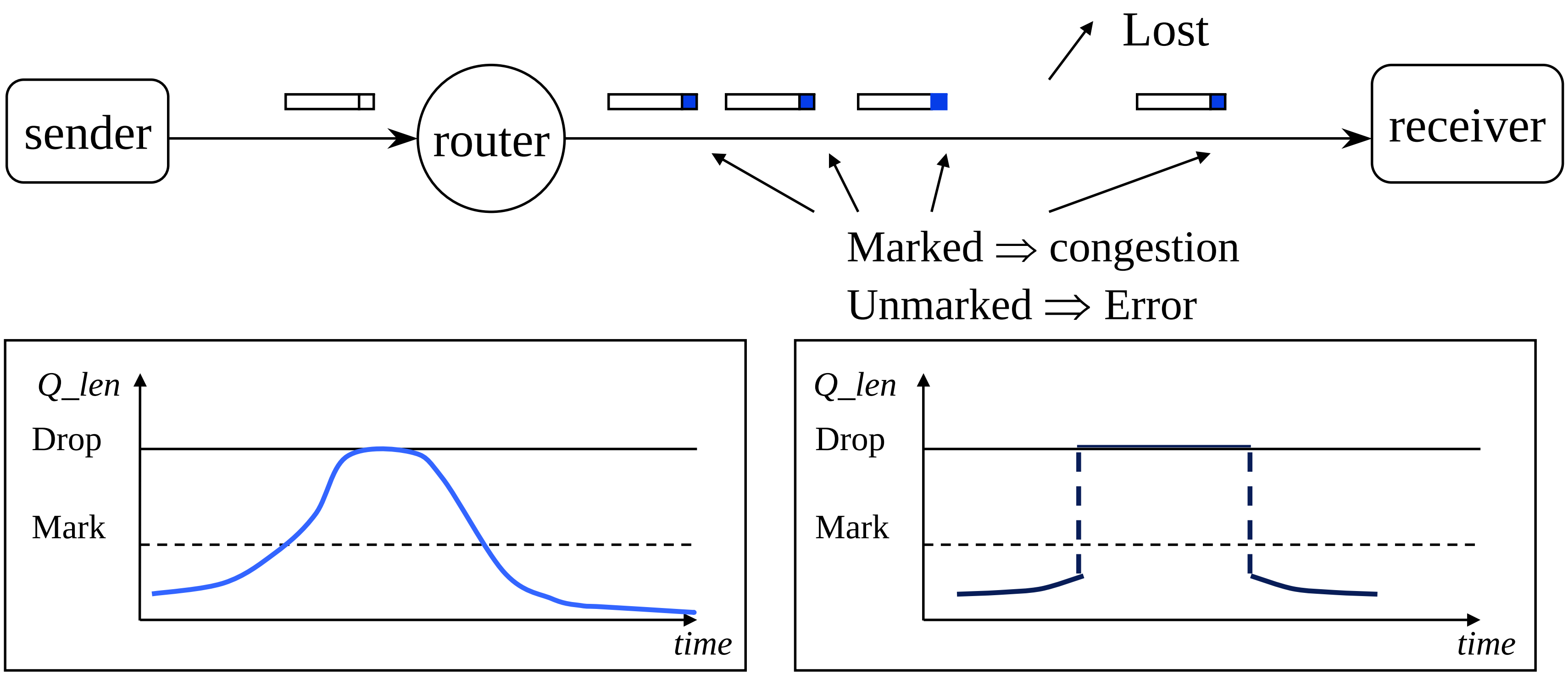
Desired Attributes of the Solution:

1. Must maintain TCP's end-to-end semantics: A packet is acked only after received by the final destination.
2. Modifications must be local: Only Base Station (BS) and Mobile Host (MH) are in the control of wireless service provider. Cannot change all locations that MH visits.
3. Must apply to two-way traffic: MH can be both a sender and a receiver.
4. Wireless links can be at the end or in the middle (satellite links)

Ref: Liu and Jain 2003

Congestion Coherence

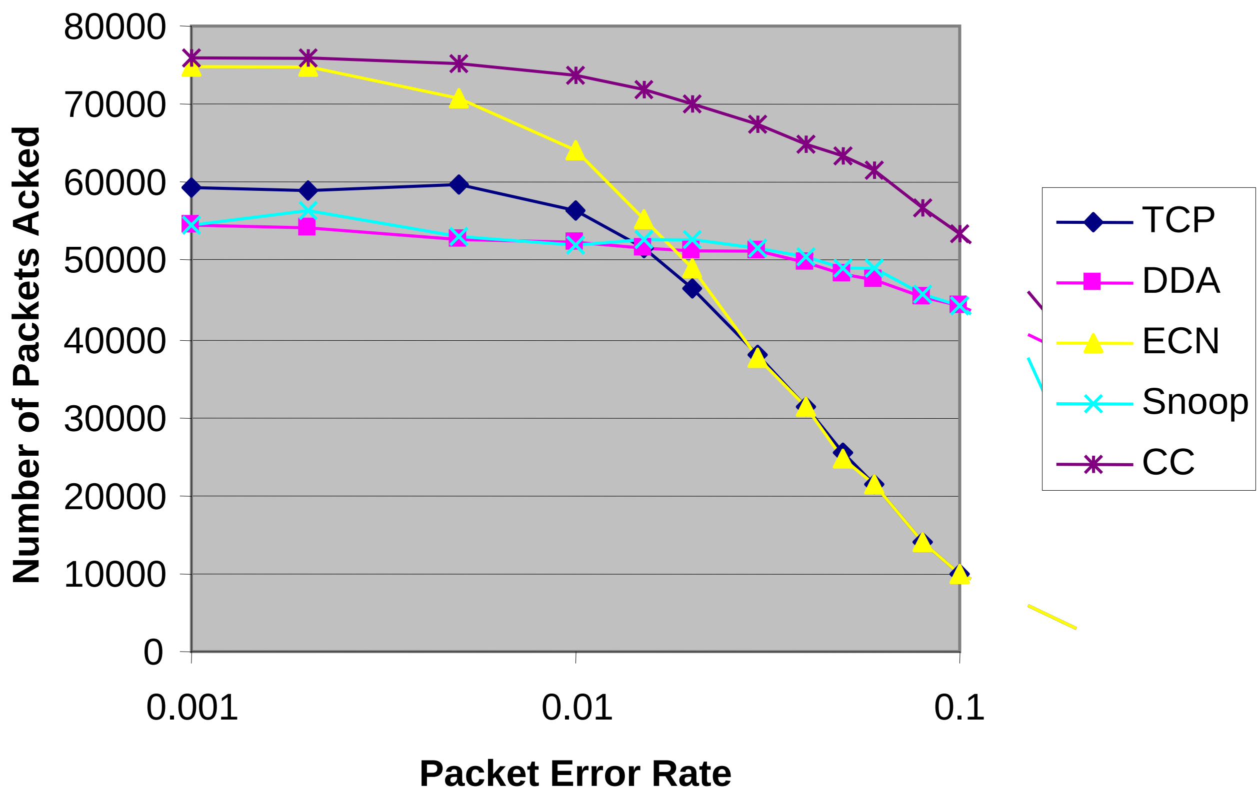
- Congestion does not happens nor disappear suddenly:
  - Before congestion reaches the point where a packet has to be dropped, some packets must have been marked.
  - After a packet is lost, some packets will be marked.



Congestion Coherence Algorithm

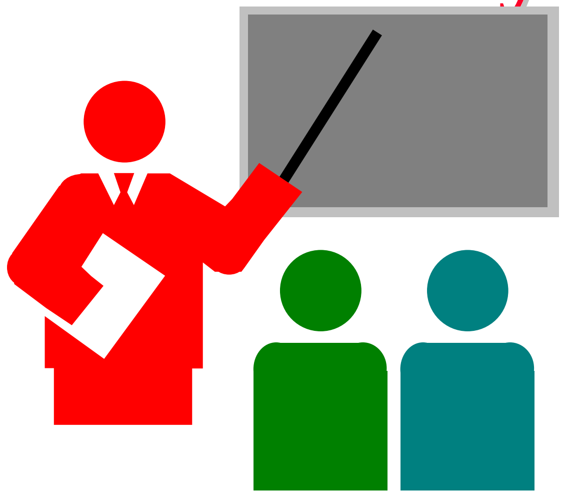
- Link layer acks and retransmissions at all wireless nodes.
- **Receiver**:
  - Out-of-order packets received check ECN bits.
  - If any packet marked, send duplicate acks Otherwise, defer the duplicate acks.
  - If expected packet arrives, drop deferred dupacks.
  - If the packet times out, release all deferred dupacks.
- **Sender**:
  - When the third duplicate acks arrives, MH checks the ECN-ECHO bits.
  - If any of thee duplicate acks carry an ECN-ECHO, MH retransmits the lost packet and reduces the window. Otherwise, TCP defers the retransmission.
  - When the expected ack arrives, cancel the deferred retransmission.
  - If the expected ack does not arrive in certain period of time then MH starts the deferred retransmission.

Goodput



- Congestion Coherence provides the highest throughput

Summary



- Frequent errors on wireless links trigger the congestion mechanism in TCP resulting in low throughput
- Key mechanisms are link level schemes to reduce/hide error losses, split TCP, TCP modification in base, receiver, or sender
- Since congestion builds up slowly, coherence of ECN bits provides a good distinction of congestion vs. errors
- On satellite links, window scaling, large initial windows, and SACK are helpful

Homework 20

- **Exercise**: A TCP entity opens a connection and uses slow start. Approximately how many round-trip times are required before TCP can send N segments.
- Hint: Write down what the CWND and total segments will be after 1 round trips, 2 round trips, 3 round trips, …

Related Wikipedia Articles

- http://en.wikipedia.org/wiki/Explicit_Congestion_Notification
- http://en.wikipedia.org/wiki/FAST_TCP
- http://en.wikipedia.org/wiki/Network_congestion_avoidance
- http://en.wikipedia.org/wiki/Slow-start
- http://en.wikipedia.org/wiki/Space_Communications_Protocol_Specifications
- http://en.wikipedia.org/wiki/TCP_Westwood_plus
- http://en.wikipedia.org/wiki/TCP_congestion_avoidance_algorithm
- http://en.wikipedia.org/wiki/TCP_window_scale_option
- http://en.wikipedia.org/wiki/Taxonomy_of_congestion_control
- http://en.wikipedia.org/wiki/Transmission_Control_Protocol
- http://en.wikipedia.org/wiki/WTCP

References

- C. Liu and R. Jain, "Approaches of Wireless TCP Enhancement and A New Proposal Based on Congestion Coherence", the 36th Hawaii International Conference on System Sciences, Quality of Service in Mobile and Wireless Network minitrack, Big Island, Hawaii, January 5-9, 2003, pp. 307a, http://www.cse.wustl.edu/~jain/papers/hicss.htm
  This paper also has references to several other papers on wireless.
- Nitin Vaidya, "TCP for Mobile and Wireless Hosts," a comprehensive tutorial presentation, 291 pp., http://www.crhc.uiuc.edu/~nhv/seminars/tcp-wireless-tutorial.ppt
- Annika Wennstr¨om, Stefan Alfredsson, and Anna Brunstrom, "TCP over Wireless Networks," http://kau.diva-portal.org/smash/get/diva2:5466/FULLTEXT01

- Read sections 9.1 through 9.6 of Murthy and Manoj

List of Acronyms

- AIMD Additive increase and multiplicative decrease
- ARQ Automatic Repeat Request
- BS Base Station
- CDMA Code Division Multiple Access
- CWND Congestion Window
- EBSN Explicit bad state notification
- ECN Explicit Congestion Notification
- ELN Explicit Loss Notification
- FEC Forward Error Correction
- FRR Fast Retransmit and Recovery
- IETF Internet Engineering Task Force
- LAN Local Area Network
- MH Mobility Header
- MSS Maximum Segement Size
- RFC Request for Comments

List of Acronyms (Cont)

RTT	Round Trip Time
SACK	Selective Acknowledgement
SCPS	Space Communications Transport Protocol
SRP	Selective Repeat Protocol
STP	Satellite Transport Protocol
TCP	Transmission Control Protocol
WTCP	Wireless Transmission Control Protocol

