TCP over Wireless

Dr. Ritika

Introduction

- Advances in wireless communication and the decreasing size and the cost of computing units have triggered an explosive growth for the mobile computing market.

- A mobile host is connected to a base station through a low bandwidth, high latency and error prone wireless channel.

- Base stations are connected to the rest of the network through the high bandwidth, low latency and relatively error free wired links.

- Bit Error Rate(BER) a wireless channel is 10^{-3} - 10^{-5}, while that of wired link is 10^{-8} - 10^{-10} or better.

Contd...

- The transmission quality is further degraded by hand-off related intermitted delay or loss of the connection as the mobile user moves from the transmission range of one base station to that of another.

- Design of TCP has to take into account the heterogeneous nature of the network.

Issues with wireless networks

- The bit error rate (BER) for a fiber optic link is usually about 10 pow -12 or better. In contrast, the hosts on wireless networks frequently move while communicating and as they share the media for communication they experience a lot of interference from the environment.

- **Link Error Rate**
- Radio transmission or infrared wave transmission Bit error rate on wireless links is found to be about 10 or worse. This mode of communication is vulnerable to interference from the environment. This high error rate poses problems for the wired networks with a wireless link .

- **Bandwidth**
- Bandwidth is a scarce resource in case of wireless networks. Compare the bandwidth of a typical Ethernet which is around 10 Mbps (100 Mbps for fast Ethernet. Bandwidth also varies highly on wireless networks. The higher layers may have to take this into consideration and use different methods (e.g. compression) to take care of this problem.

Issues with wireless networks

- **Mobility**
- During movement the data sent to the wireless host is lost. TCP at the destination interprets this loss as congestion and invokes congestion control mechanisms, which is unnecessary, as when the move is complete the wireless host will start receiving data again. This causes the performance of TCP to degrade. There also might be degradation of performance due to frequent recalculation of routes to the moving wireless host.

- For a better throughput, a segment loss due to wireless link errors must be detected and retransmitted as quickly as possible.

Protocols

- Many approaches have been proposed to improve the performance of TCP over networks with wireless links. These have been broadly categorized as Transport Level Proposals and Link Level Proposals.
- Various Protocols considered
  - Wireless unaware TCP at the fixed host
    - I-TCP (Indirect TCP)
    - Berkeley Snoop Module
    - M-TCP
    - Delayed duplicate acknowledgements
  - Wireless aware TCP at the fixed host
    - Fast Retransmit
    - Mobile TCP
    - Multiple Acknowledgements
    - Discriminating congestion losses from wireless losses
    - Distinguish losses by making two connections

Wireless unaware TCP at the fixed host

- The TCP sender is unaware of the losses due to wireless link so the TCP at the sender need not be changed.
- The non-congestion related losses are hidden from the TCP at the fixed host (sender) and hence the TCP at the fixed host remains unmodified.
- This approach is based on the intuition that as the problem is local it should be solved locally and the TCP should be independent of the behavior of the individual links.

Indirect TCP

I-TCP overcomes by utilizing the resources of Mobile support router(MSR's) to provide transport layer communication between mobile hosts and fixed hosts.

I-TCP is fully compatible with TCP/IP on fixed network.

Any interaction between Mobile Host(MH)& Fixed Host (FH) are split into two separate interactions.

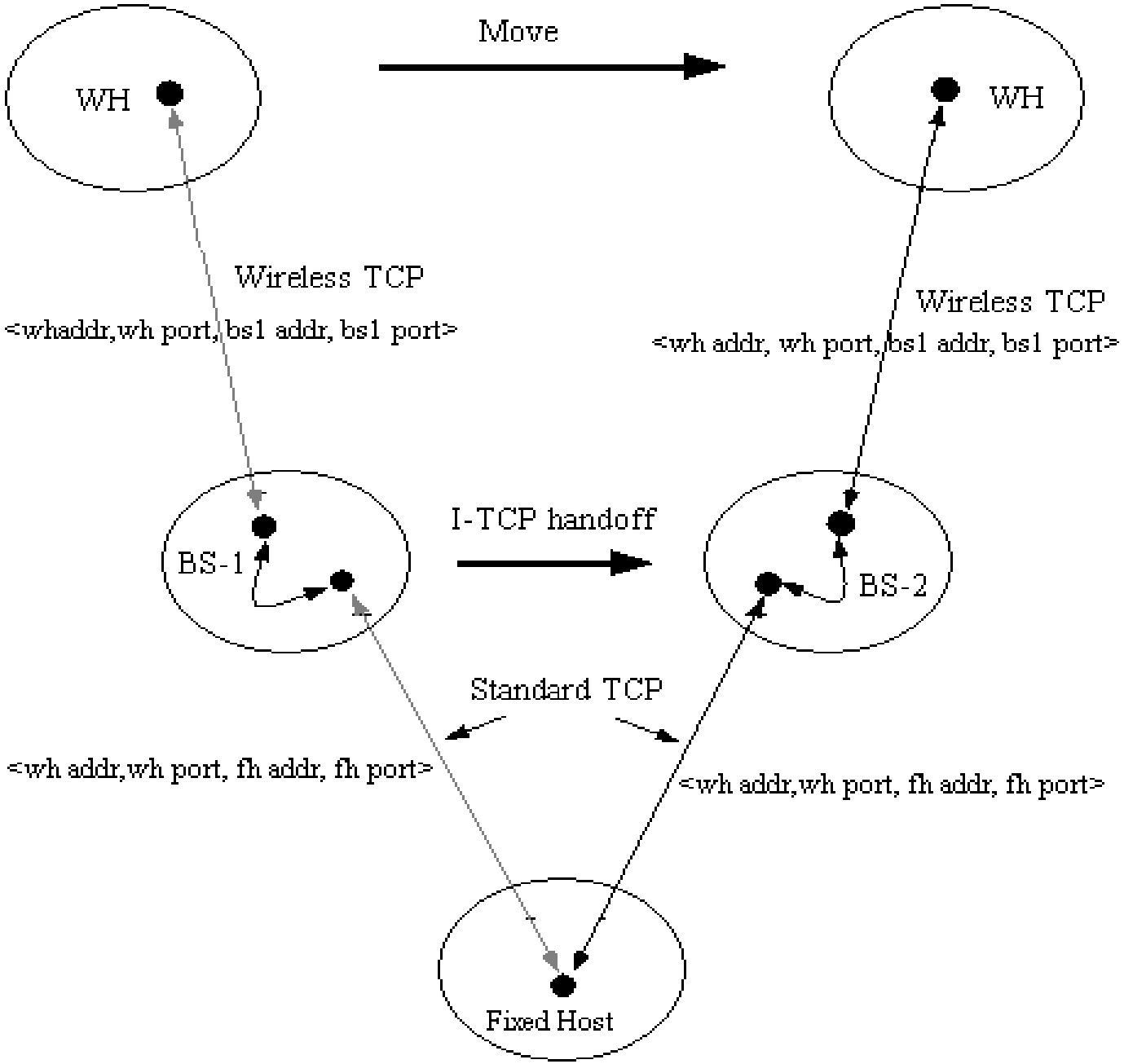
- Connection between wireless host MH and fixed host FH goes through base station BS

- FH-MH = FH-BS + BS-MH

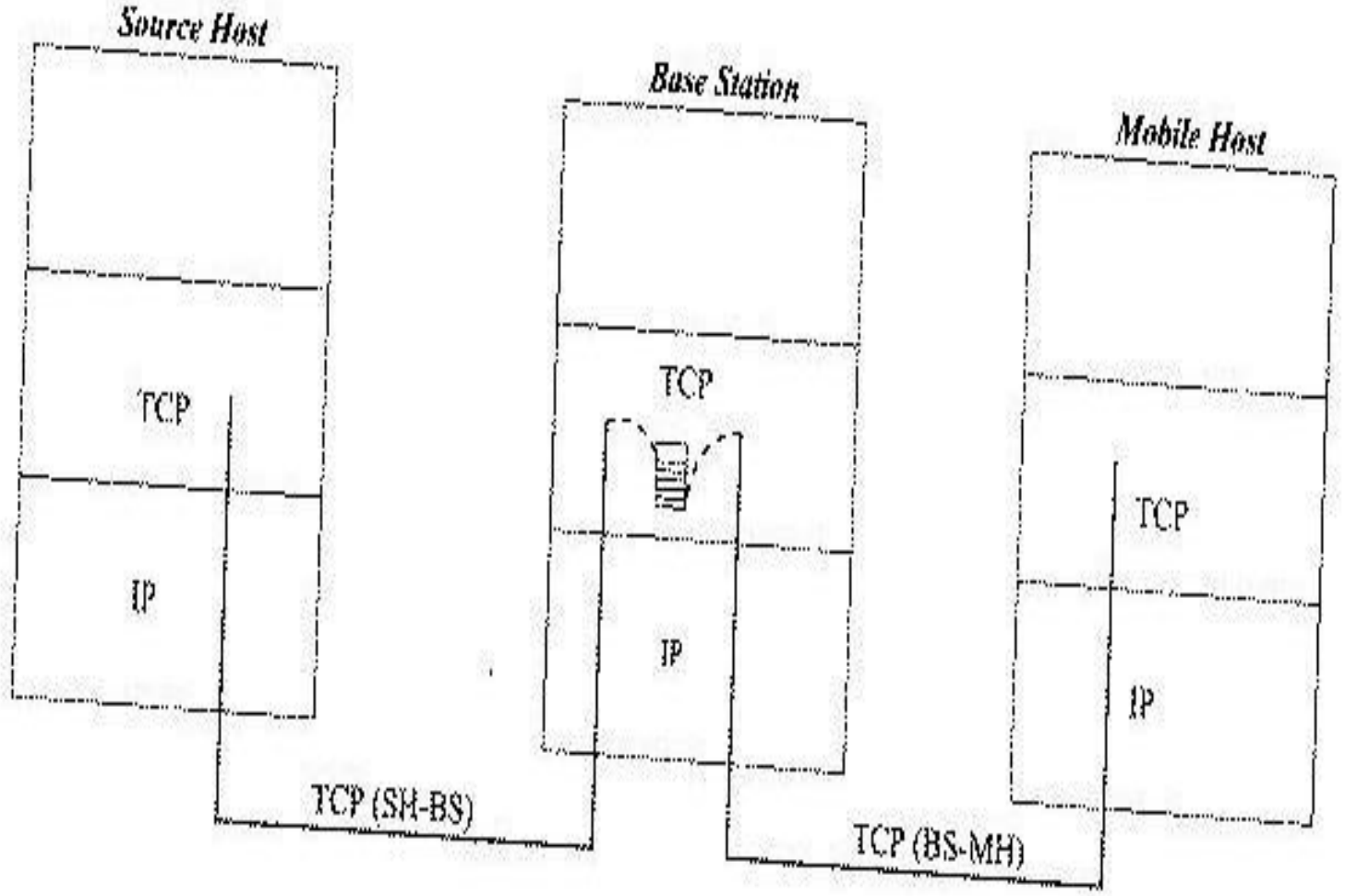
Using I-TCP MH sends a request to current Mobile support router(MSR) to establish a connection.

Indirect TCP

- It uses special I-TCP calls, instead of regular system calls.
- If the MH switches cells during the life time of an I-TCP connection the center point of the connection moves to new MSR.
- When the MH switches cells, the states associated with two sockets is handed over.
- The FH is completely unaware of the indirection.
- The connection is not re-established at the new MSR.


Move
WH
WH
Wireless TCP
<whaddr,wh port, bs1 addr, bs1 port>
Wireless TCP
<wh addr, wh port, bs1 addr, bs1 port>
I-TCP handoff
BS-1
BS-2
Standard TCP
<wh addr,wh port, fh addr, fh port>
<wh addr,wh port, fh addr, fh port>
Fixed Host

ITCP Architecture



Limitations of I-TCP

- The drawbacks of this method are:
- 1) If there are frequent handoffs then the overhead related to the connection state transfer between the base stations may be large and add delays.
- 2) The base stations have to be complex and with large buffers in case of heavy traffic.
- 3) I-TCP violates the semantics of end to end reliability.

Snoop Protocol

- Proposed by Elan Amir, Hari Balakrishnan, Srinivasan Seshnan & Randy H. Katz - University of California, Berkeley.

- Improves on Split connection approach by retaining the end to end reliability semantics.

- Wireless errors are handled by local retransmission at the base station.

- Reduces interference between TCP retransmission and link level retransmission.

- The probability of time out at the TCP sender is assumed small.

- Snoop uses features like link level retransmission time out.

Snoop Protocol

In this approach the network layer software is not modified anywhere in the wired network except at the base station.

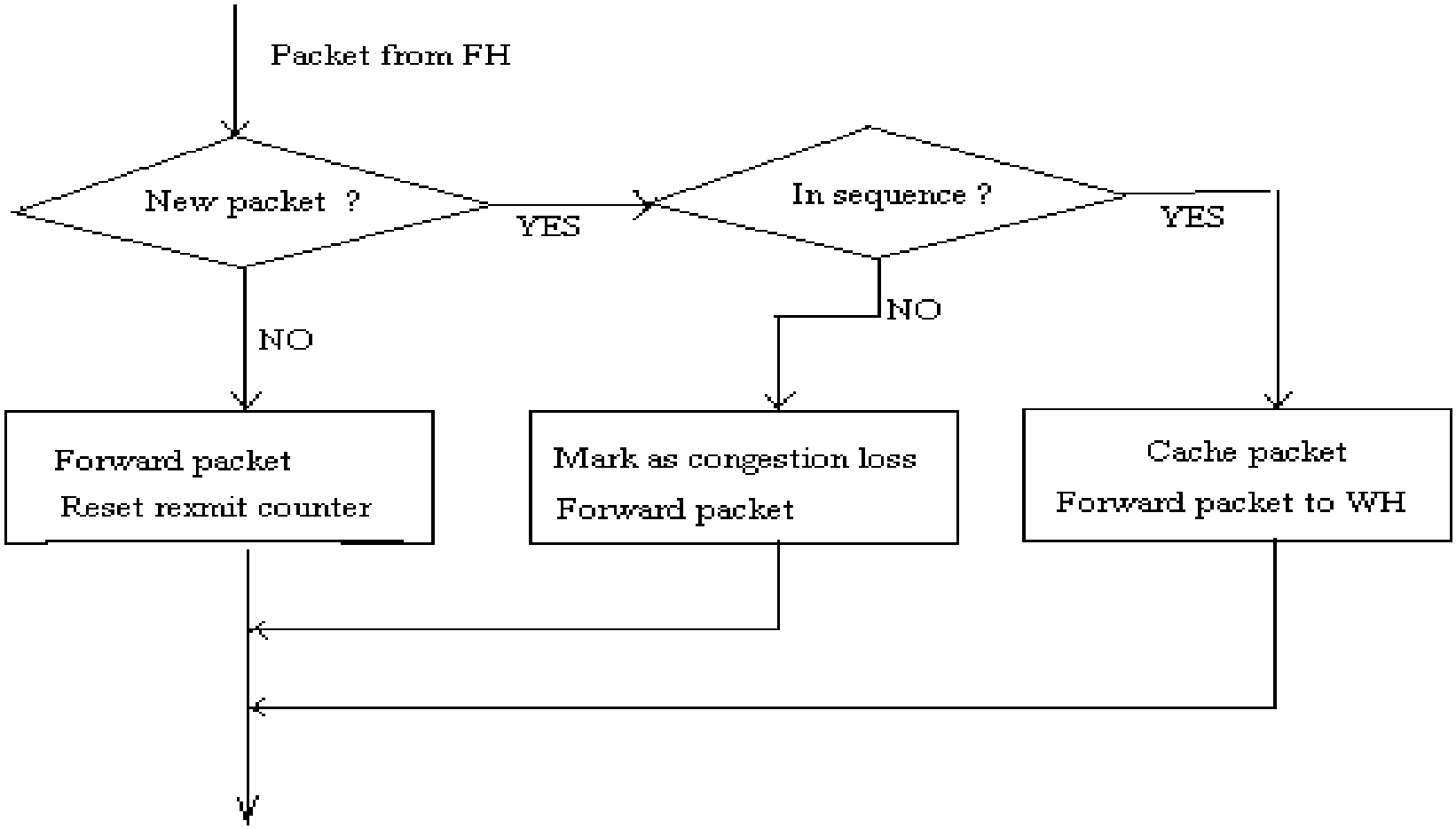
The router code at the base station is modified to cache data meant for the wireless host.

A layer called, snoop layer is added the base station. This layer looks at every packet on the connection in either direction.

The module caches the packets that are sent by the fixed host to the wireless host but have not yet been acknowledged by the wireless host.

Actions at the base station when it receives a data from the fixed host are shown next slide.

Actions at base station on receiving data from fixed host



Snoop_output()

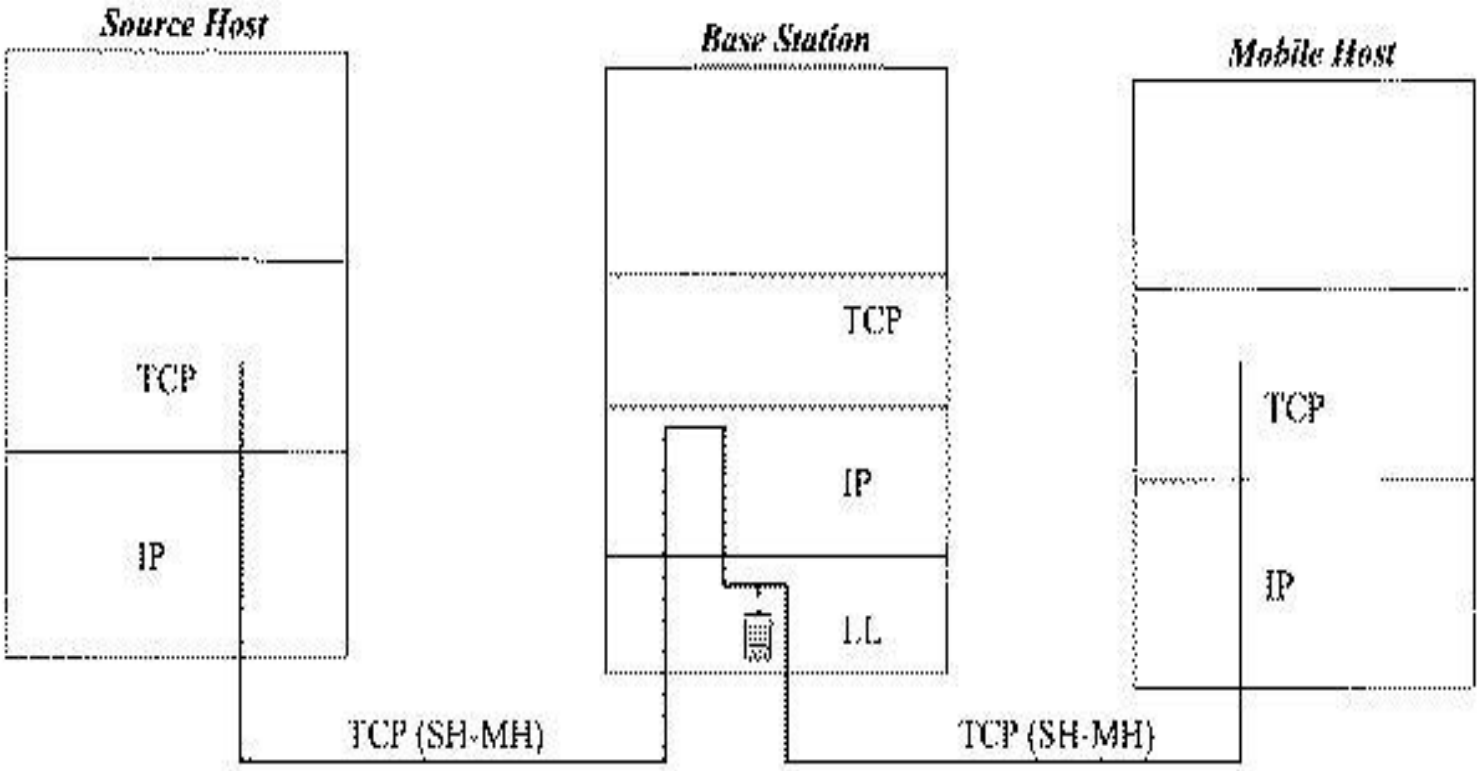
Types of packets from the fixed host

A new packet in the normal TCP sequence.

An out of sequence packet that has been cached earlier.

An out of sequence packet that has not been cached earlier.

Snoop Architecture



On receiving a packet from the fixed host the module stores the packet in its cache and then passes it to the wireless host.

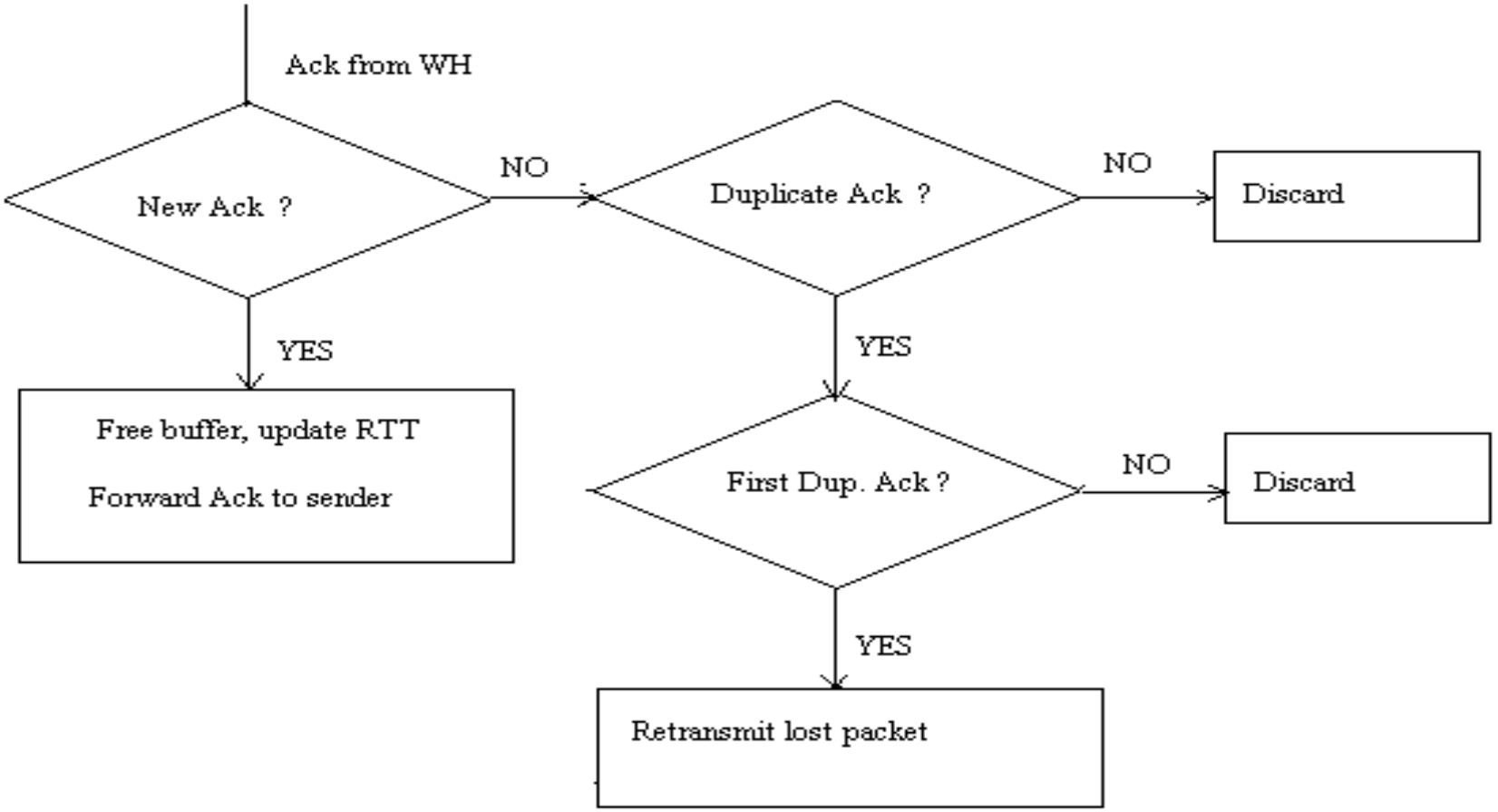
If the packets are lost on the wireless link then the base station gets repeated acknowledgements for the lost segment from the wireless host.

The snoop module on detecting this loss checks if it has the packet in the cache and retransmits the packet at the same time suppresses the ACK to the fixed host, otherwise forwards the ACK to the fixed host and lets the sender recover from the loss.

Types of ACK's from the Mobile host

- *A spurious ACK*
- *A DUPACK*
- *A New ACK*
- *Time Out*

Actions at base station on receiving an ACK from wireless host.



It also identifies a group of base stations where the wireless host can move, and broadcasts the packets meant for the wireless station to this group. This avoids expensive state transfer during the handoff. When the wireless host moves to the new base station in the group it already has the cache of packets so the transport state gets recovered quickly.

M-TCP

- A new architecture is proposed for modern cellular networks to support high bandwidth multimedia services and real time audio and video .

- It has three layers. Mobile Hosts are at the lowest level and communicate with MSS (Mobile Switching Station or base station) nodes in each cell. A supervisory host controls several base stations (SH).

- SH is connected to a wired network and handles most of the routing and other protocol details. It also maintains connections and handles flow control.

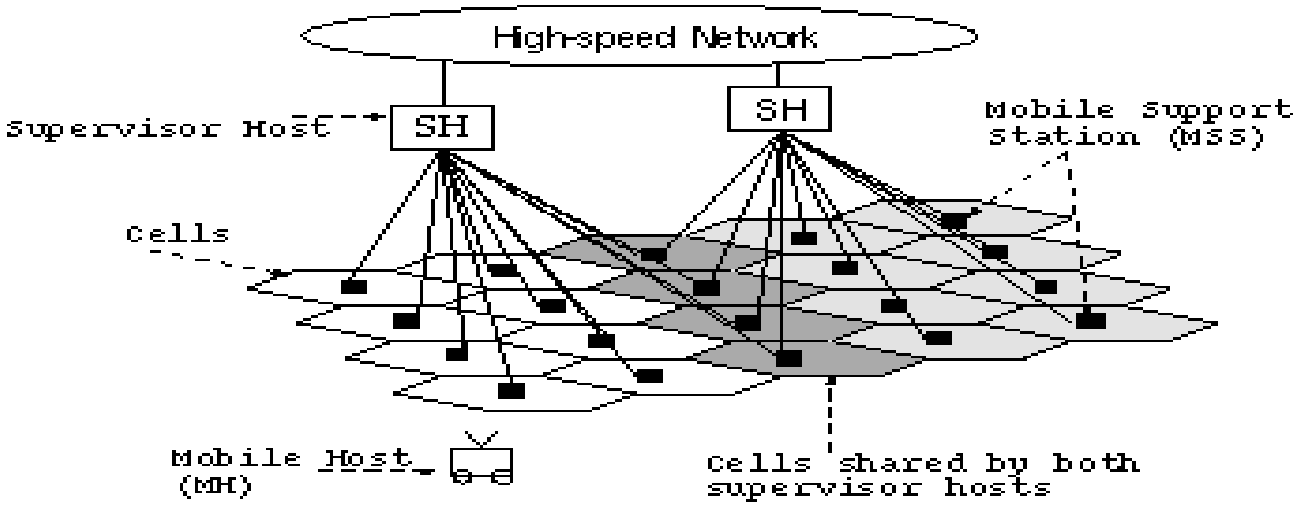
- When a wireless host moves from one cell into another, two base stations need not perform state transfer if the same SH controls them.

M-TCP

This also uses the split connection method similar to I-TCP. The connection between the fixed host and the wireless host is broken up into two parts.

The split connection method performs better than standard TCP because the base station (where the connection is split) is usually at one hop away from the wireless host and hence provides better control over the losses and also can adapt quickly to the dynamic mobile environment because of low RTT.

The base station sender makes its congestion window small in case of losses. The window builds up fast when new acknowledgements arrive from wireless host because of small RTT.



Architecture for M-TCP

M-TCP

- TCP at the sender is not modified whereas the TCP at the base station is called SH-TCP (Supervisory Host TCP). The base station and the wireless host communicate using M-TCP.

- When the base station receives the data from the sender it forwards it to the wireless host but defers the acknowledgement (for the segment) to the sender until it receives an acknowledgement from the wireless host.

- M-TCP can be designed so that it is optimized for wireless links. Data compression can be used by M-TCP to make efficient use of wireless bandwidth. However the complexity at the base station is very high.

M-TCP

- If the size of advertised window at SH-TCP is W and it has received acknowledgement from wireless host for w<=W then the SH-TCP acknowledges w-1 bytes and defers ACK for the last byte.

- Now at this point if the wireless host undergoes a handoff or a period of data loss the base station sends the deferred acknowledgement and advertises the window size of zero, this causes the sender to go in a persist state. In this state all timers are frozen and does not close its congestion windows and does not back off its timers.

- When the wireless host regains the connection it sends a greeting packet to the SH. SH_TCP send a duplicate acknowledgement with the window update packet to the sender so that it can resume transmitting data. This method provides a solution to the problem of periodic disconnection.

MTCP

Similar to I-TCP except that the last TCP byte of data is acknowledged to the source only after it is received by the mobile host.

Source falsely believes that every byte of data except the last byte is received by the receiver, it can take remedial measures based on whether the last byte is received or not.

Delayed Duplicate Acknowledgements

- This scheme attempts to imitate the behavior Snoop scheme .

- However in snoop scheme we needed to make modifications to the TCP whereas this solution is 'TCP Unaware' i.e. no changed in the TCP are needed.

- The base station does not need to look at the TCP header so this method may be specially useful if the transmission is encrypted (the snoop approach fails in case of encrypted transmission).

- The base station needs to implement a mechanism to perform link level retransmissions.

- This method tries to reduce the interference between TCP retransmissions and the link level retransmissions.

Delayed Duplicate Acknowledgements

- When the TCP receiver receives out-of-order packets, it sends the duplicate ACKs for first two out-of-order packets but if it gets more of them then it defers the ACKs for these packets for a time period of, say t.

- If during this time period t, it gets the next in-sequence packet then it discards the duplicate ACKs.

- If it does not get the in sequence packet during this time it sends all the deferred duplicate ACKs. on the wired network.

Delayed Duplicate Acknowledgements

This method performs better in a scenario when the packet losses are due to wireless transmission errors and performs bad when the losses are due to congestion.

The overall performance depends on the relative frequency of the two types of losses:

If the packet was lost due to transmission error on wireless link and if the time t was chosen large enough for the link level retransmission to take place then the TCP receiver gets the packet before t and hence does not send the third duplicate ACK and hence the sender does not fast-retransmit.

If the packet was lost due to congestion on the wired link, TCP receiver will delay the third duplicate ACK for time t. This can make this method perform worse than the standard TCP.

Wireless Aware TCP at the fixed host

In this approach the fixed host (sender) is aware of the existence of the wireless link in the network and is able to distinguish the losses due to transmission error on wireless link from those due to congestion.

The sender can avoid invoking congestion control algorithms when the losses are due to the wireless link.

Fast Retransmit

- Addresses the issue of mobile host hand off.
- During mobile host hand off TCP segments can be lost or delayed.
- The Mobile host has to unnecessarily wait for a long duration for the source to retransmit.

Fast Retransmit

This method presents a solution to the problem where only one or two segments are lost.

When a wireless host moves from one cell to other cell some packets are lost. The sender waits till a timeout occurs and treats this as congestion and reduces the window and retransmits.

This solution overcomes this problem by making the wireless host resume the communication immediately after the handoff without waiting for the timeout at the sender.

The wireless host sends a threshold number of duplicate ACKs to the sender. This prompts the TCP at the sender to reduce the window size to half and begin retransmissions.

Mobile TCP

The delay characteristics when a wireless host switches to a different network is different from when it moves from one cell to another in the same network and the data loss from these two reasons is different from data loss due to congestion in the wired network.

The fixed host interprets these packet losses due to handoffs or interface switching as congestion and invokes the congestion control methods including reducing window size. This is not desirable.

Mobile lets the base station tell the sender whether the loss is due to handoff in the same network or if it is due to interface switching.

Mobile TCP

The sender then marks the packets and retransmits them once the mobile host has completed handoff.

- In case of interface switch the wireless host enters a new network which may not have the same network characteristics as the previous one. So when the TCP at the sender knows about the interface switching it resets window size, congestion window threshold size(ssthresh), estimates of RTT (Round Trip time) and RTO (Retransmission Time out) values and begins slow start.

- But if the wireless host has moved to a cell in the same network then the values of window size and ssthresh are halved and the RTT value remains the same.

Multiple Acknowledgements

- This method distinguishes the losses due to congestion or other errors on the wired link and those on the wireless link. This method uses two types of ACKs to isolate the wireless host and the fixed network.
  - **ACKp:** This partial acknowledgement carries a sequence number S. This tells the sender that packets till S-1 have been received by the base station.
  - **ACKc:** This is the normal ACK in TCP.
- Sender uses standard TCP with slow start, congestion avoidance, fast retransmit and fast recovery.

Multiple Acknowledgements

- Let RTT be the end-to-end round trip time.
- RTT (w) be the round trip time between the base station and the wireless host. The sender retransmits after waiting RTO time for end-to-end ACK.
- RTO(w) is the maximum time between a packet is received at the base station and the time when it got a acknowledgement from the wireless host.

Multiple Acknowledgements

- When a base station gets a packet from sender it starts a RTO (w) timer and sends it to the wireless host. If the timer goes off the base station sends a ACKp to the sender.

- If the base station receives an out of order packet which is not cached then it will try to deliver this packet to the wireless host.

- If the packet cannot be delivered for RTO(w) time and if all earlier packets have been received by the base station then it sends a ACKp to the sender.

- The ACKs from the wireless host are forwarded only if they are needed by the TCP at sender and the ACKs that trigger unnecessary retransmissions are discarded.

Multiple Acknowledgements

When a sender gets a ACKp from a base station the sender understands that base station has received packets upto S-1 but is having difficulty sending these to the wireless host i.e. it has not received the acknowledgement from the wireless host even after waiting for RTO(w) time.

So the sender updates RTO to avoid end-to-end retransmissions. It marks the packets corresponding to this ACKp.

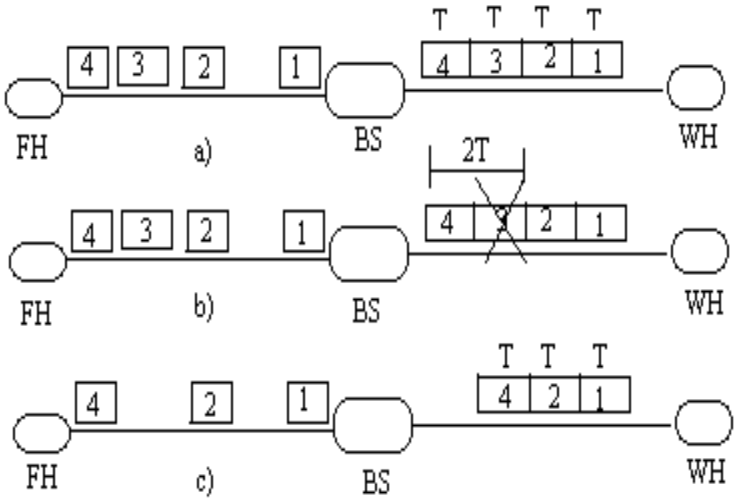
On a timeout the sender checks if the packets are marked if they are not it acts as normal TCP would.

If they are marked then it will not retransmit, and will not do any congestion control methods except backing off the timer.

Discriminating congestion losses from wireless losses

- This method again uses the model where the network is wired and the last hop is wireless.
- It is also assumed that the wireless link is the bottleneck in the network. As the wireless link is the bottleneck the packets arrive at the base station early and they are queued at the base station.
- So most packets are transmitted back to back on the wireless link.
- The sender performs bulk transfers.
- The packet inter-arrival time is defined as the time between arrival of consecutive packets.

Discriminating congestion losses from wireless losses



In case 'a' no packets are lost so packet inter-arrival time is same as time required to transfer the packet on the wireless network (T).

In case b the packet 3 is lost so the time of arrival of 4 after 2 is 2T as packet 3 was on lost on the network before using T on the link.

In case 'c' packet 3 is lost due to congestion so the inter-arrival time between packet 2 and packet 4 is T.

Discriminating congestion losses from wireless losses

This method can be useful to distinguish losses due to congestion and losses due to transmission errors on the wireless link and the TCP at the ends can be designed to take advantage of this knowledge and improve performance.

Let **Tmin** be the minimum inter-arrival time observed so far.

P0 denotes the out of order packet received by the receiver.

Pi was the last in-sequence packet received before P0.

Tg is the time between arrivals of P0 and Pi

n be the number of packets missing between Pi and Po.

If **(n+1)Tmin <= Tg <=(n+2)Tmin**

n missing packets are assumed to be lost due to wireless transmission errors otherwise they are assumed to be lost due to congestion.

Distinguish losses by making two connections

- The basic model for this method is that the network is wired and the last hop from the base station to the wireless host is wireless.
- The key idea in this algorithm is that the sender distinguishes between congestion on the wired network and the transmission errors on the wireless part and the wireless host assures that if the packets are dropped due to transmission error the sender retransmits them before its timeout.
- When a fixed host wants to communicate with a wireless host it opens two connections one with the base station and other with the wireless host. The connection between the sender and the base station is called control connection(estimate the congestion on the wired link).
- The packets on these two connections are expected to be routed in the same way and hence are affected same by the congestion.

Distinguish losses by making two Connections (Cont..)

- Sender sends packets on the control connection in regular intervals sufficiently spaced so as not to cause overhead.

- The sender periodically compares the fraction of acknowledged packets on the two connections and checks if the packets are lost due to congestion or due to transmission error on the wireless link.

- If the acknowledged fraction is significantly different in these two cases then it concludes that the error in the wireless link is causing the packets to be dropped so the sender does not apply congestion control method and does not reduce the window size and continues the increment as before.

Distinguish losses by making two Connections (Cont..)

If the two fractions are same then the congestion control methods are applied as in normal TCP.

When the packets are lost the TCP at the sender has to wait for timeout after which it retransmits the packets.

When the wireless host learns about this loss it sends duplicate ACKs.

The sender on seeing the duplicates ACKs knows that its previous packets have been lost and immediately resends the packets without waiting for the timeout.

TCP extensions for satellite networks

Satellites are very useful for communication because they can reach geographically remote areas and places which lack terrestrial communication infrastructure(useful for mobile users). The satellite networks can be classified in following types:

- *Asymmetric satellite networks:* This is characterized by large data rate in one direction and very less in the other. These satellite systems are unidirectional and use non-satellite path in other direction(for ACKs). This is seen to be a problem for TCP.

- *Satellite link with last hop:* These provide service directly to the users. This can provide shared high speed downlink to users with low speed. This causes the network to be asymmetric as the return path may be terrestial path.

TCP extensions for satellite networks

- *Hybrid satellite networks :* Satellite links are located at any point in a network and act as a link between two routers.

- *Point-to point Satellite networks:* This is a pure satellite network with all hops going through satellites.

Issues

- All satellite networks are inherently characterized by high delays because of finite speed of light and the altitude of the communication satellites. GEO long distance causes the ground-to-satellite-to-ground propagation delay to be 239.6 milliseconds for a radio signal. For ground stations it turns out that the delay is 279ms. So the round trip delays for a message and reply could be as high as 558ms.

- It can be more than this if there are multiple satellite links. Sometime low earth orbit satellites(LEO) or medium earth orbit satellites(MEO) are used. These require a constellations of satellites so that they are constantly visible to the ground station, The propagation delay for LEO may be around 80ms.

- The radio signal attenuates inversely in proportion of the distance. The large distance of satellites makes the signal weak This causes the BER to be high. Another issue with satellite networks is of limited radio spectrum and there is a restricted amount of bandwidth available which is controlled by licenses.

Issues

TCP cannot make efficient use of satellite network because of he following characteristics of satellite networks.

*Large delay*bandwidth product:* The delay bandwidth can be defines as amount of data on the link at any time to utilize the available channel bandwidth. This poses a problem for TCP.

Long feedback loop: Because of the long delay on the satellite links, it may take a long time for the sender to know if the packet was received at the destination. This is also bad for interactive applications like telnet and for congestion control algorithms.

Transmission errors: Satellites have high BER and the TCP at the sender cannot determine if the packets are dropped due to congestion or due to transmission errors. It interprets this as congestion and reduces the window size which worsens the situation.

Issues

- *Variable round trip delays:* In low earth orbit constellations the delays varies from time to time. This may or may not affect TCP performance.

- *Handoffs:* In non-geosynchronous satellites TCP connections need to be transferred from one satellite to other or from one ground station to another. This handoff affects the TCP performance adversely.

- *Asymmetric use:* Some connections use satellite link for one direction and terrestial link for opposite direction this poses problem for TCP.

Proposed Solutions: Slow start and congestion avoidance

Slow start algorithm wastes a lot of bandwidth on satellite networks.

Because of large delay*bandwidth product it takes a large time to increase the congestion window to fill the link and hence effectively utilize the bandwidth.

Delayed ACKs also cause wasted bandwidth during this slow start phase.

One method to deal with this is to increase the initial value of congestion window. It is suggested that following value of initial window yields good results

Min(4*MSS , max(2*MSS , 4380))

This large window causes large packets to be sent in first round trip time so there will be more ACKs and the window size can be increased fast.

To deal with the problem of delayed ACKs, 2 segments are sent initially instead of one so there is no need to wait for the delayed ACK timer.

Fast Retransmit

- TCP uses timeouts to detect lost segments.
- When the timer expires the TCP retransmits the data and performs congestion control by setting ssthresh is set to half the value of the window, the window size is halved. The window size now is increased by one segment for each duplicate ACK it receives.
- This algorithm uses three duplicate ACKs to trigger retransmission of the lost segments of data and the sender can retransmit them without waiting for the timeout. This is called fast retransmit.
- Now it adjusts the window sizes which is called fast recovery.

Fast Retransmit (cont..)

- Fast retransmit increases performance however in some cases it is seen to degrade performance.

- In connections with large congestion windows. New segments are introduced during the recovery which can trigger multiple fast retransmits. This can reduce the congestion window multiple times for one loss event.

Large window sizes

TCP should be extended for larger windows. This help improve performance of TCP on satellites.

- The standard TCP allows a maximum window size of 65535 bytes. TCP throughput is limited by:

  throughput = window size / RTT

- Geosynchronous satellite have a channel RTT of 560 ms. The maximum throughput that can be achieved by standard TCP over satellites is limited to:

  throughput = 65,535 bytes / 560 ms = 117,027 bytes/second

Selective Acknowledgements

When multiple packets of data are lost, the TCP at sender waits for the timeout and determines which segments have to be retransmitted.

As there are no ACKs coming in (which are used to clock new segments in network) the sender invokes slow start and restarts transmission which is very time consuming and degrades the performance of TCP.

TCP receivers can inform senders which packets have arrived using Selective acknowledgments (SACKs). This way TCP recovers more quickly form the lost segments and avoids unnecessary retransmissions.

With SACKs, the sender can determine which segments need to be retransmitted in the first RTT after loss detection.

This way the sender can send segments at a appropriate rate, sustain the ACK clock and avoids the slow start too.

TCP for Transaction

- TCP uses three way handshake to establish a connection. This requires 1-1.5 round trip times and this time can be high.

- This method (T/TCP) eliminates this startup time. The two hosts need to set up connection only once. Next time T/TCP can bypass this three way handshake and the sender can transmit the data in the first segment itself with the SYN.

- This can enhance performance in case of short request response type of connections between the sender and the receiver.

- This method requires changes at both sender and receiver side.

Explicit congestion notification

In this approach the intermediate node in the network (routers) inform TCP senders about congestion in the network by using Explicit Congestion Notification (ECN). This can be classified in two parts:

- **Backward Explicit Congestion Notification (BECN):** The router sends messages directly to the sender and informs it about the congestion. This can be done by using ICMP Source quench message. The sender on getting this message reduces its window size (it reduces its rate of data transfer).

- **Forward Explicit Congestion Notification (FECN):** In this approach when the routers see the congestion they mark the packets with a special tag before forwarding the packet. The data receiver on receiving this marked packet sends this congestion information to the sender in the ACK packet. The sender then invokes the congestion control methods.