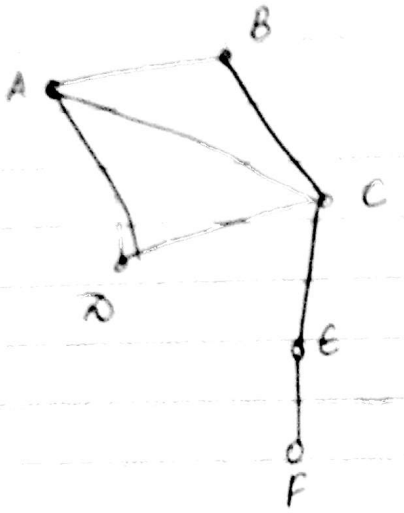
WK 21 145

25 WEDNESDAY MAY

14/10/2016

Intra-Domain Routing

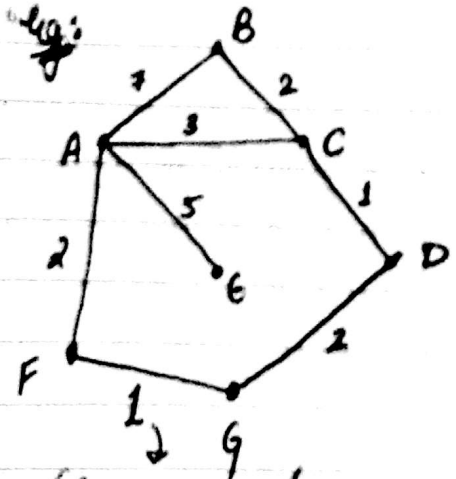
* **Distance-Vector:** Every node shares with neighbours: its distance to all nodes
  ⊗ Count to ∞ problem



Forwarding Table:

Dest	Next hop	cost

* **Link-State Routing:** Every node shares with all nodes its distance to its neighbours.
  - All nodes have entire network graph
  - Dijkstra algorithm runs locally by all nodes



Link weight/cost

DV
⊗ tells G & C

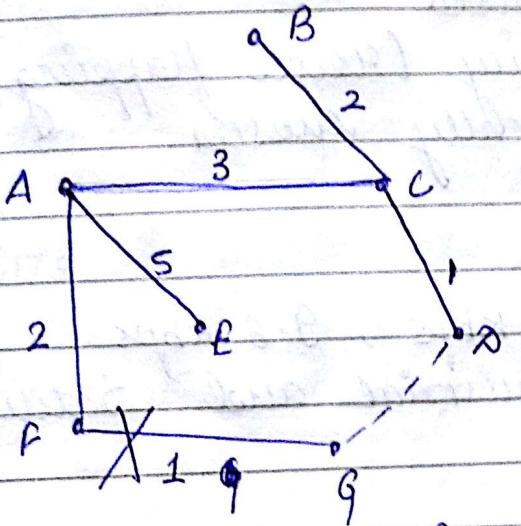
$\begin{pmatrix} A,4 \\ B,3 \\ C,1 \\ E,9 \\ F,3 \\ G,2 \end{pmatrix}$

LSR
⊗ broadcasts to all

$\begin{pmatrix} C,1 \\ G,2 \end{pmatrix}$

2016

The shortest path tree for the above network, from A to other nodes :



Suppose F-G link fails

In LSR, F & G send a broadcast that F-G has failed.
⇒ All rerun Djikstra & Very fast convergence.

Advantages of LSR over DV

- Recovers quickly from link failures
- No count to ∞ problem.

Disadvantages of LSR over DV

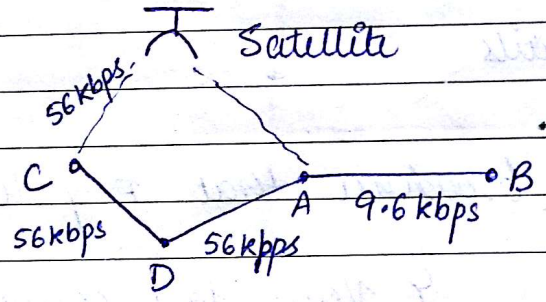
- More information transmitted to nodes.

Q. How to set link weights ?

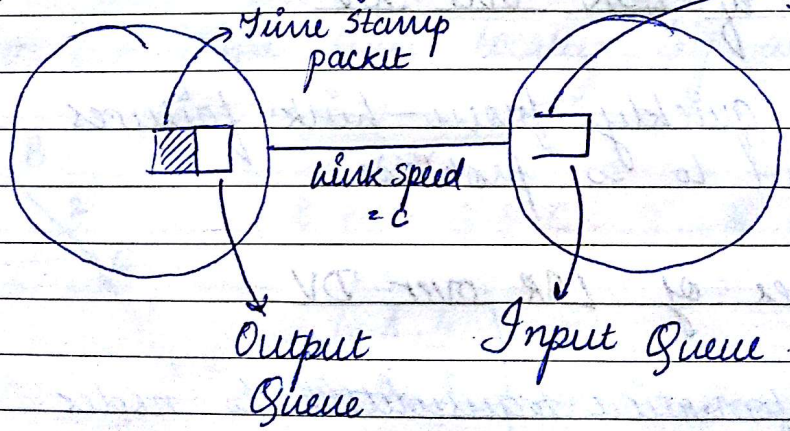
- Want network utilised
- Can have instability (route flapping) if weights dynamically chosen.

ARPANET :

2 link speeds : 56 kbps, 9.6 kbps.
2 Types of links : Terrestrial and Satellite



Initially, use latency on links



Difference = Queuing delay + Speed of light prop. + Store & forward delay
↓
for each packet

Take average over time window to get link weight.

Idea: Penalise links which are congested.

Let the link speed be 'C' (BITS/SEC) and pkt. size be 'P' (BITS).

BITS in queue ahead of pkt. = Q.

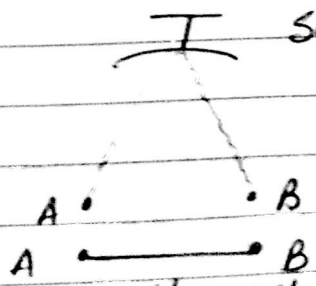
$$\text{Diff} = Q/C + P_{prop.} + P/C$$

→ heavily dependent on C

If C = 9.6 kbps, P/C is going to be very large for the given Queuing delay, Q.
⇒ Generally, C = 56 kbps is preferred.

$P_{prop.}$ → Propagation delay is going to depend on the 2 links.

Eg:



Satellite link is much longer
⇒ Penalised a lot in Prop. component.

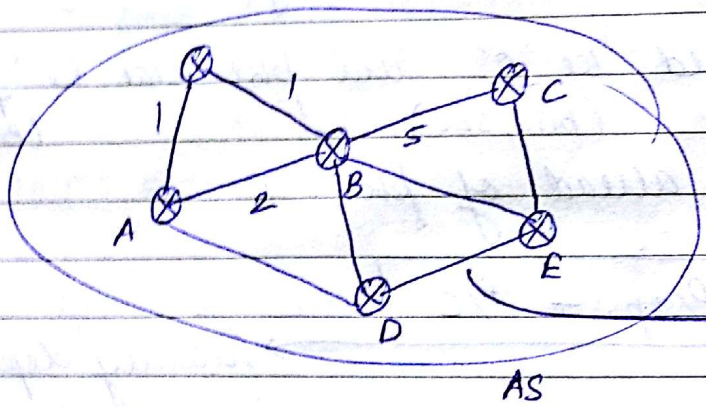
Sunday 29

* Route Swapping Problem.
A link might look congested at one time
⇒ Switching to another idle link
⇒ It might get congested
⇒ Switch again.

18/10/2016

Intra Domain Routing

Distance Vector / Link State Routing



→ How to set link weight / cost ?

A-B-C
A-D-E-C

↪ Flip : Frequency change in weights.

↓

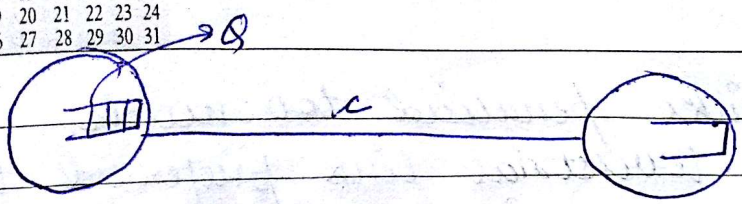
Can lead to loops in shortest path algorithms has not finished running at some node.

ARPANET :

Types of links : Satellite, Terrestrial
Speeds : 9.6 kbps, 56 kbps.

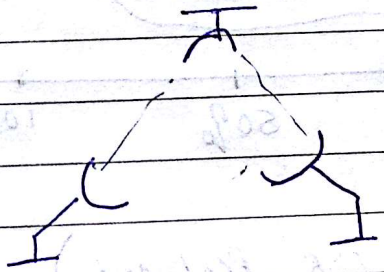
A-B-C 1, 2, 3, 8, 9
A-D-E-C 4, 5, 6, 7
At C, 1, 2, 3, 4, 5, 8, 9, 6, 7
↪ Pkt reordering (We dont want this)

Want :
- Higher Speed → lower weight
- Congested links → Higher weight
- Chosen routes to be stable



Initially, used latency on link as wt. latency for a packet = Q/C + PROP + P/C

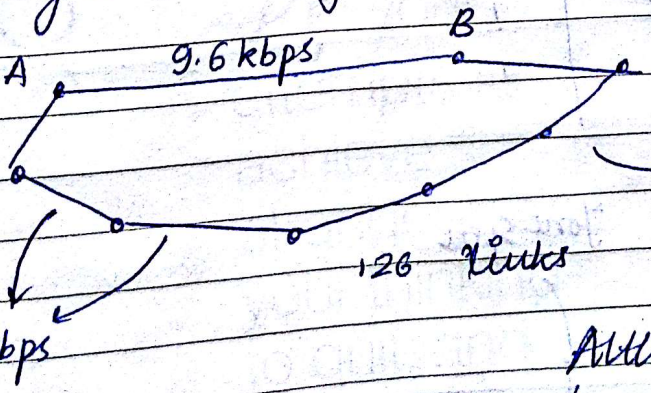
Q = Queue Size in bits
C = Link Speed
P = Pkt. size.



Problems:

(a) Route Oscillations
- Queuing delay increases weights of congested links leading to new shortest paths.

(b) Range of link weights is very large.



Routing protocol was preferring this path b/w A & B and not direct. Although, more data can be pushed, but it is a sheer waste of resources.

Penalising long paths?

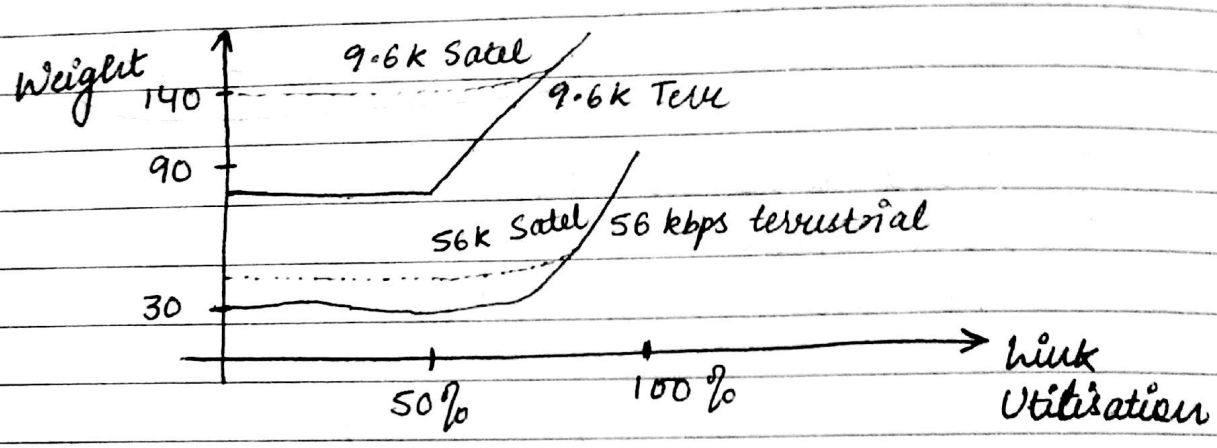
May'16

M	T	W	T	F	S	S
30	31					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

June'16

M	T	W	T	F	S	S
		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			

@ Satellite links penalised too much
- 9.6 kbps terrestrial link preferred to 56 kbps SATEL. link.

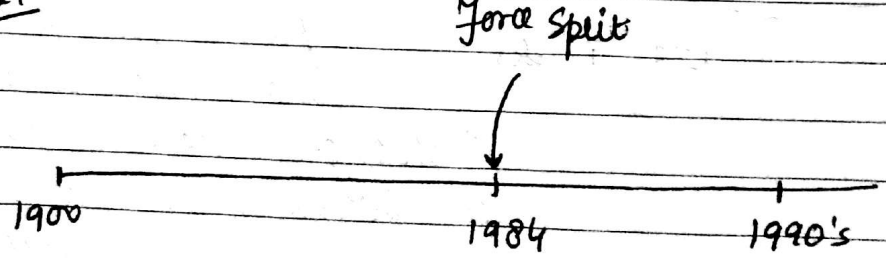


OSPF (a LSR Protocol)

Link weight (Default) = $\max\left(\frac{10^8}{\text{link Speed in Bits/sec}}\right)$

link Speed	Wt.
10 Mbps	10
100 Mbps	1
1 Gbps	1

AT&T

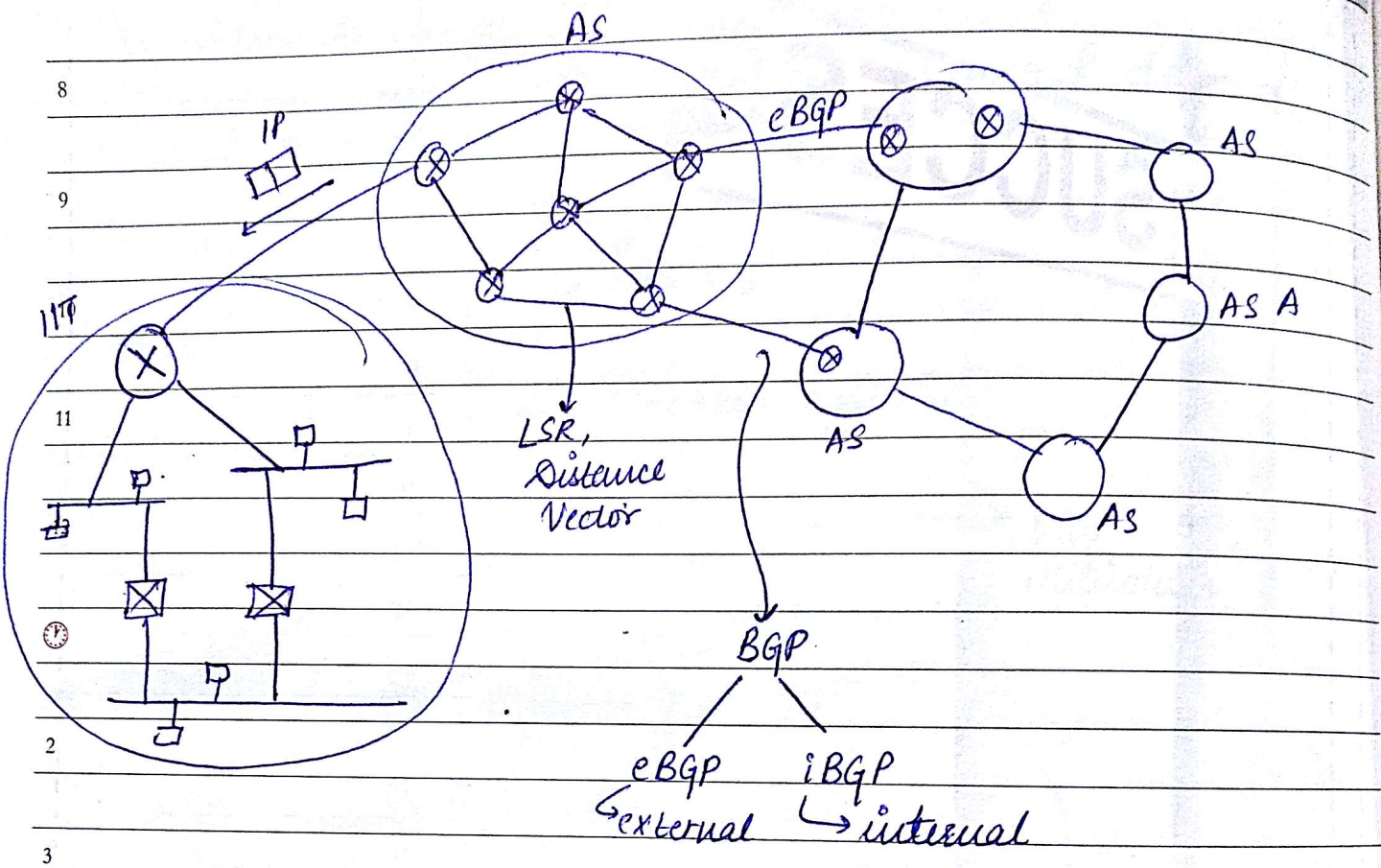


Bell Labs

AT&T → Lucent (Bell)

AT&T (BELL)

AT&T → BABY (BELL CORE) MESSAGES BELLS'

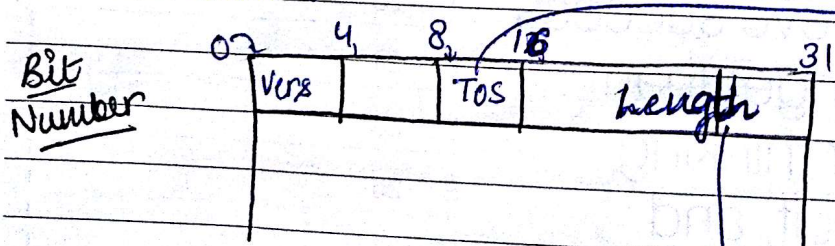


Routing:

19/10/2016

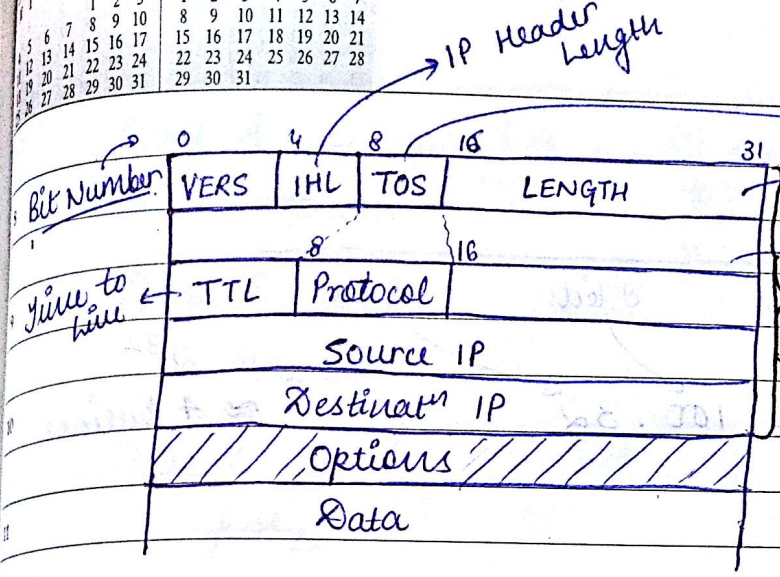
Physical Header	DLL Header	IP HDR	Transport HDR	.....

IPv4 Header:



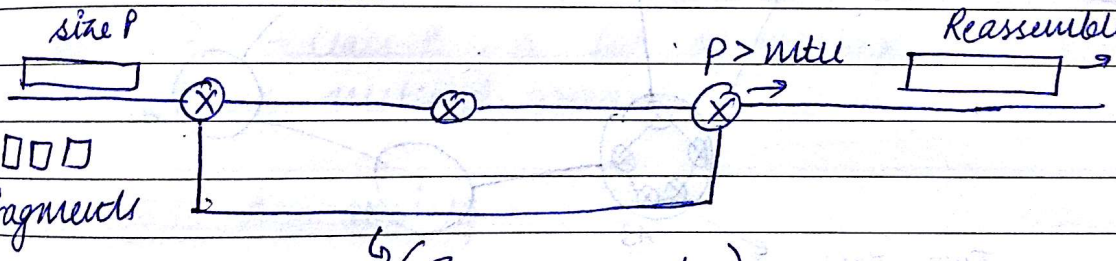
→ Type of Service:
→ Can indicate different QOS for pkt.
→ Not used much.

→ Length of IP Datagram



* **Fragmentat" / Reassemble :**

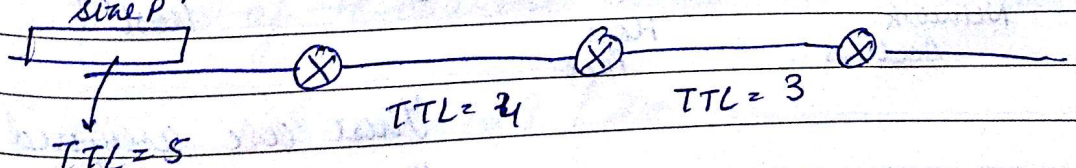
eg: LTE
VoLTE → Voice
File Transfer.



↳ (Suppose mtu)
↳ Max. Transmission Unit

* **Time To Live (TTL) :**

TTL decremented by 1 at every IP router & Checksum Updated at every hop.
If TTL = 0, discard packet.



* **Protocol :**

Identifies Transport Protocol
[eg: TCP(6), UDP(17), ICMP(1)]

WK - 22 156-210

04 SATURDAY JUNE

May'16

M	T	W	T	F	S	S
30	31					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

June'16

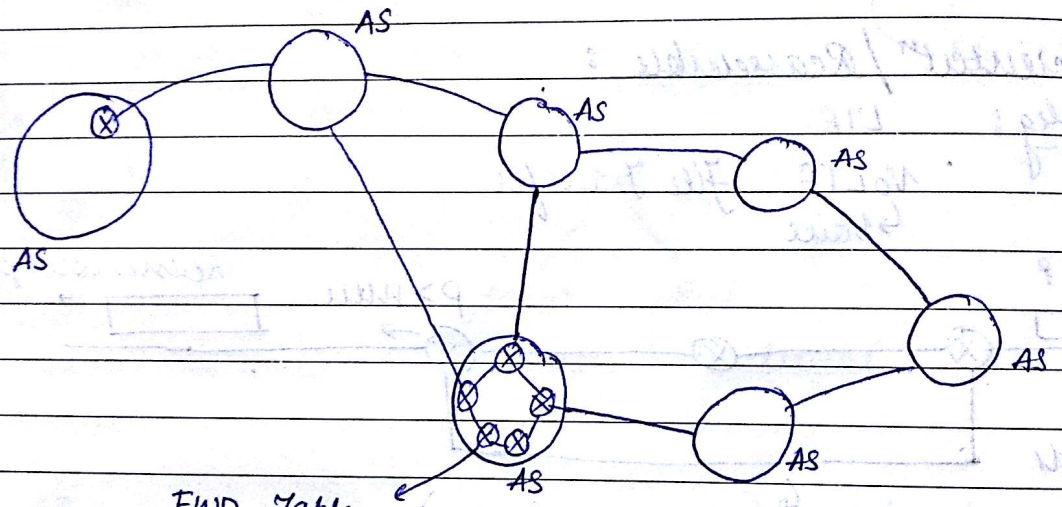
M	T	W	T	F	S	S
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30			

IP Address:

8 bits | 8 bits | 8 bits | 8 bits

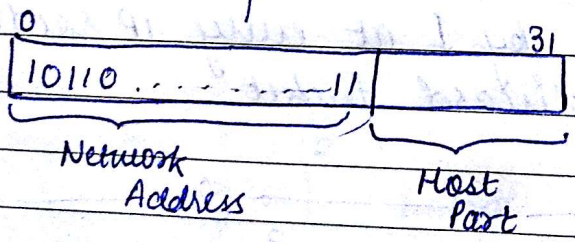
Human Readable form: 73 . 12 . 100 . 32 (decimal)

$2^{32} \approx 4$ billion



FWD Table

Dest.	Next hop



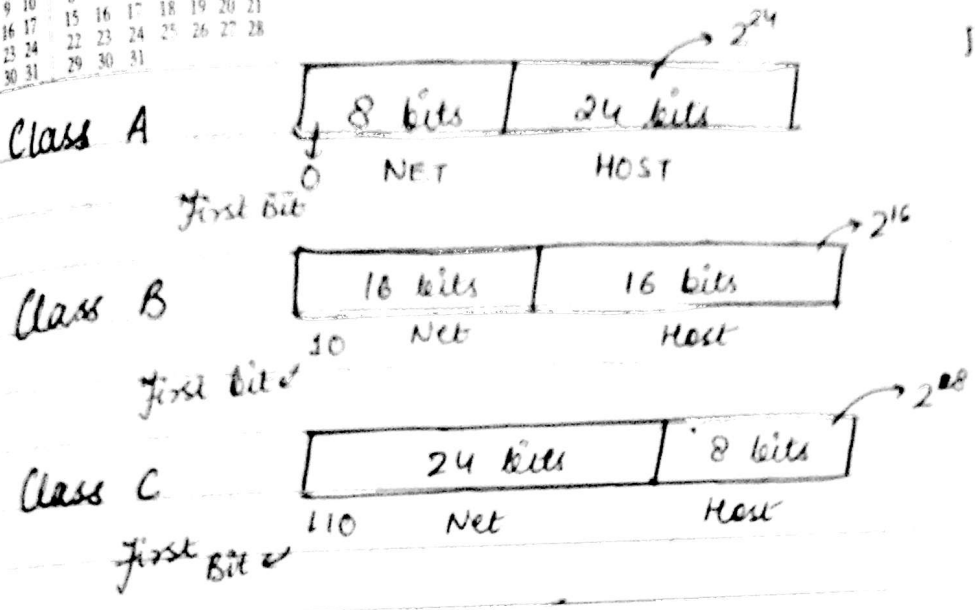
} IP Address of each host.

These are assigned such that hosts from an autonomous system get a block of IP addresses, with same most significant bits

05 Sunday

2016

MESSAGES



Class D, E

↳ There maybe a lot of waste of addressing, say host address ~ 280 Since C → $2^8 \cdot 256$, we might go for 2^{16} → Class B, a lot of wastage 9 bits
⇒ Better method needed.

CLASSLESS ADDRESSING

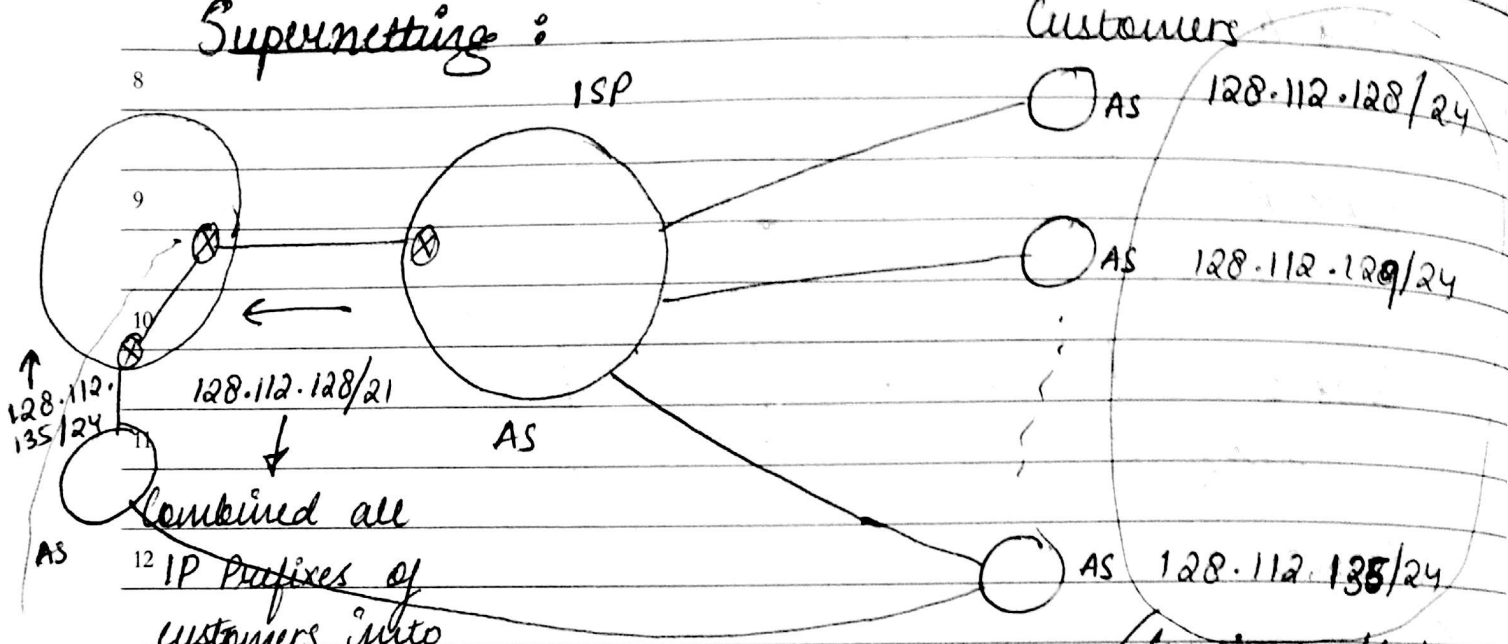
$a.b.c.d/n$ → Notation
→ # bits for network address.

Note: You need to specify atleast 24 bits a.b.c, to get a network address. It will then give a list of IP addresses under that network address.

eg: 128.112.128/24

→ Unlike classes which has n= 8, 16, 24 /8, /16, /24, we have /n to have independence of length.

Supernetting :



Combined all IP Prefixes of customers into single prefix

Dest. IP Prefix	Next hop
128.112.128/21	...
128.112.135/24	....

Longest Prefix Match

Find the longest prefix match.

Customers

AS 128.112.128/24

AS 128.112.129/24

AS 128.112.135/24

Contiguous blocks ⇓ Can be clubbed.

8 prefixes are contiguous

/21 & 24 in all.)

3 bits for 8 contiguous address past.

(Rest is common)

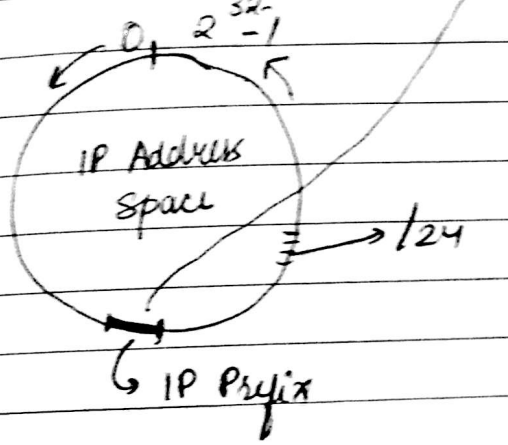
21/10/2016

IP

a.b.c.d/n

First n bits

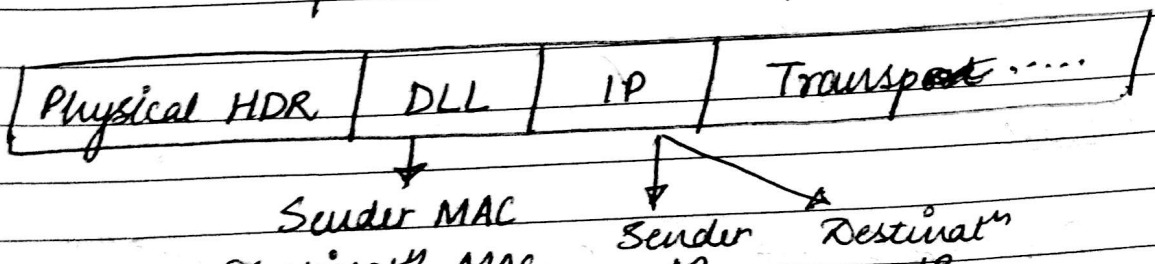
Eg: 73.0.0.0/8

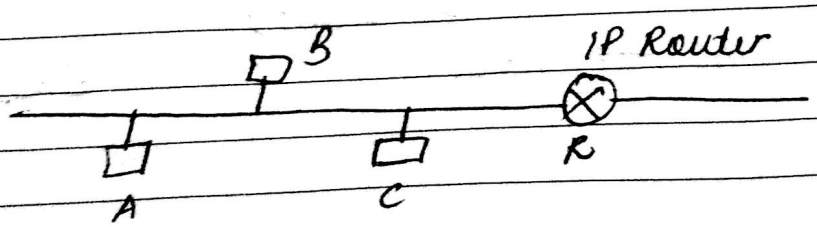


Supernetting: Combines many IP prefixes representing contiguous blocks in IP address space into a single prefix.

ARP: Address Resolution Protocol

Get MAC address, Given IP Address.





WK-23 161-205

09 THURSDAY JUNE

May'16

M	T	W	T	F	S	S
30	31					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

Examples:

1) A has IP of B, but not its MAC address
2) Router 'R' gets packet with destination IP that of B, but R doesn't know MAC address of 'B'.

Network	ARP
DLL	
Physical	

ARP Frame:

Preamble	Dst MAC	SRC MAC	Type/Length	ARP Packet	CRC
8 bytes	6 bytes	6 bytes	2 bytes		

- Type/Length → Value > 1536 → Indicates type of frame
- ARP Packet:
  - A_{MAC} : Sender MAC
  - A_{IP} : Sender IP
  - All zeroes : Target MAC
  - B_{IP} : Target IP
- For IP_{v4} address to Ethernet MAC.
- For Broadcast, set DST. MAC to all ones.

ARP Request:

DST. MAC → All Ones (Broadcast)

SRC. MAC → A_{MAC}

A → A_{IP}, A_{MAC}

B → B_{IP}, B_{MAC}

ARP Pkt :
- A_{MAC} : Sender MAC
- A_{IP} : Sender IP
- All zeroes : Target MAC
- B_{IP} : Target IP

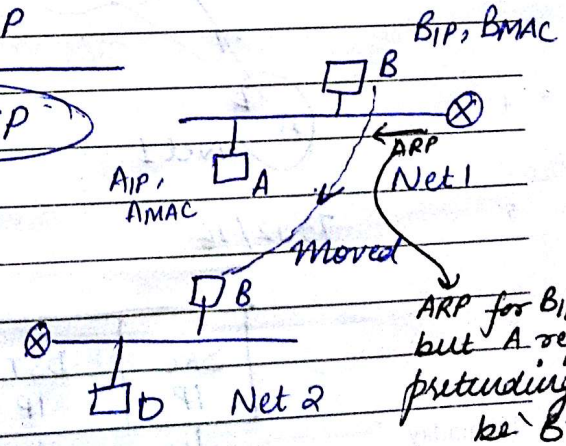
All nodes received frame since target IP matches B_{IP}, B replies.

B replies : Unicast ~~from~~ frame to A

- **ARP Reply :**

B_{MAC} : Sender MAC
B_{IP} : Sender IP
A_{MAC} : Target MAC
A_{IP} : Target IP

Mobile IP



* If someone wants to be malicious as someone else using the Mobile IP
Designed to address cases where some can move.

* If B sets A as its "home agent", while B is out of Net1. Say a pkt for B arrives when B is out, A acts as B (pretends) and replies on behalf of B.

WK-23 163-203

11 SATURDAY JUNE

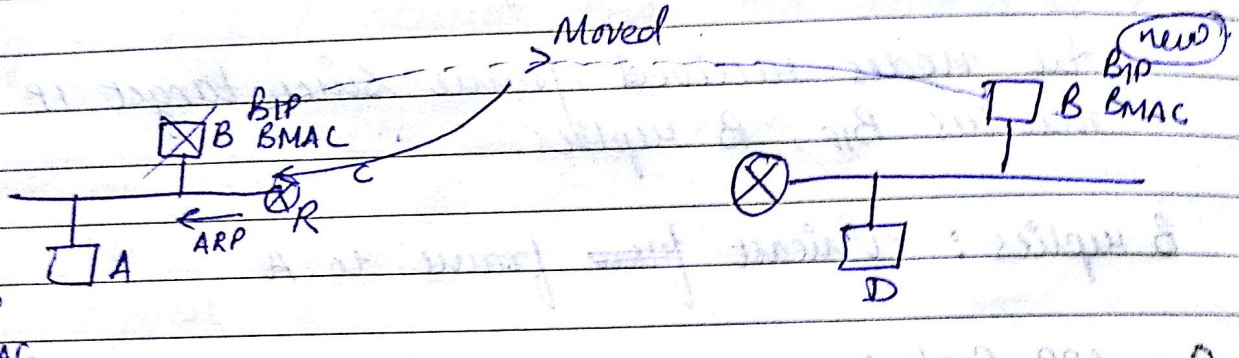
May'16

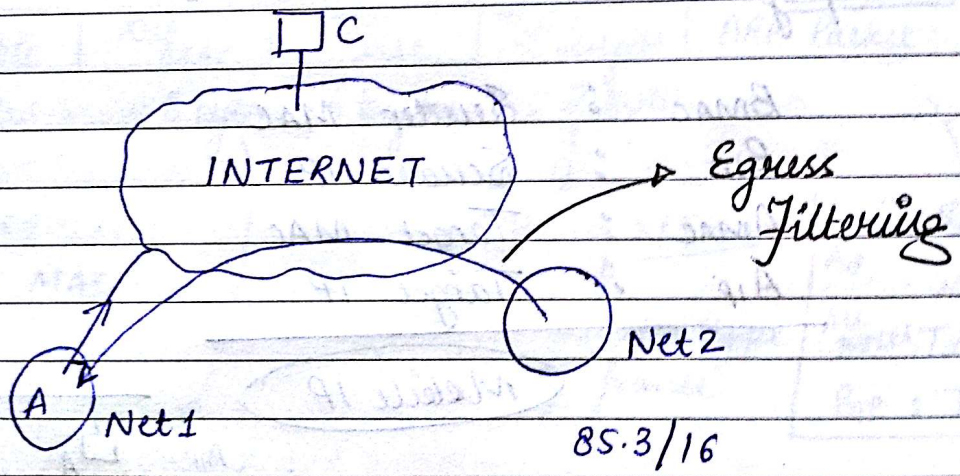
M	T	W	T	F	S	S
30	31					1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29

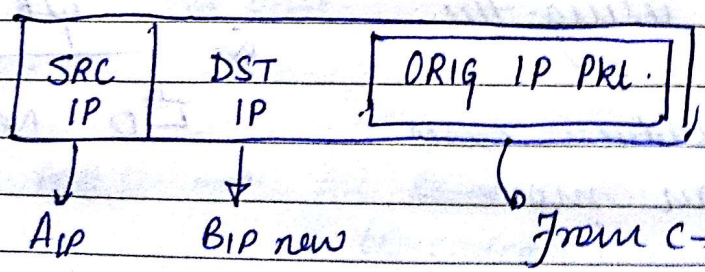
June'16

M	T	W	T	F	S	S
		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			

* All packets received by A for B are forwarded to B_{IP}^{new}.
B will have to configure its Mobile IP to set A as home agent.
Mobile IP Protocol decides who the home agent is.







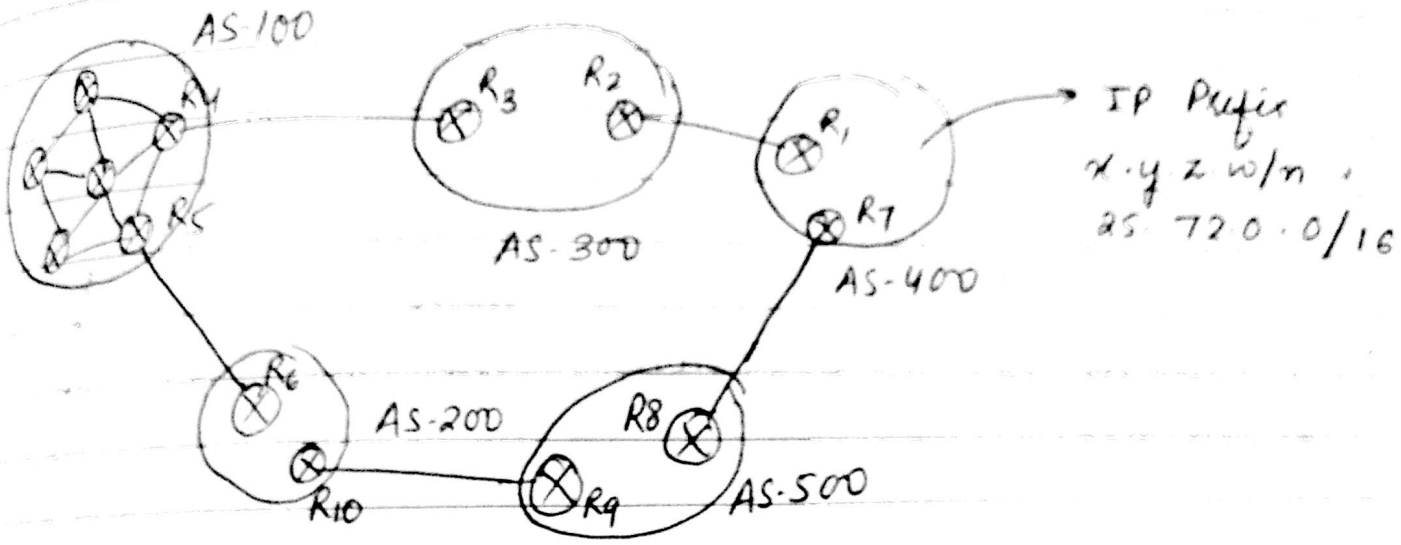
If B at Net2 sends a packet, it will get blocked.
So, B encapsulates the packet and sends it to A.
A then sends it to C.

12 Sunday

MESSAGES

Border Gateway Protocol (BGP)

01/11/16



IGP → Interior Gateway Protocol.

- BGP Advertisement

IP Prefix (R1 (in)) — BGP Attributes (R2 (in))

Example:

- AS 400 advertises to AS-300
  25.72.0.0/16 — AS-400 (AS Path) — other attributes (optional).
- R3 (in AS-300) tells R4 (AS 100)
  25.72/16 — (AS 300, AS-400) AS Path
- R5 tells R6
  25.72/16 — (AS-100, AS 300, AS 400) AS Path

Rule: If an AS advertises a prefix, then it guarantees that it will forward packet matching prefix along AS-path advertised.

× ×

▶ R_7 tells R_8 25.72/16 AS-400 (Path)

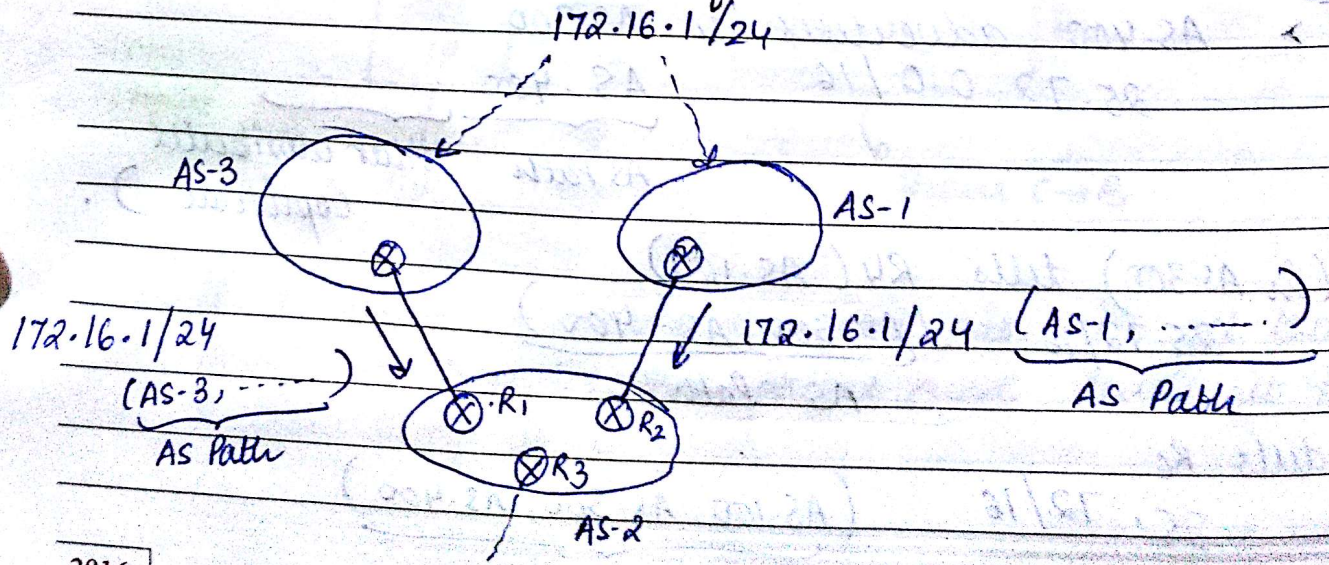
▶ R_9 need not advertise/announce this info to R_{10} of AS-200.

Q At any point, if I have 2 paths to reach an AS (point), which one should I choose?

⇒ Many factors contribute into the decision making like the AS-path and all the other attributes.

BGP Attributes

1) LOCAL-PREF {Local Preference}



AS 2 has the freedom to choose a path.

- R_1 tells other BGP routers (R2, R3, ...) within AS2.
  172.16.1/24 (AS-3, ...) local Pref = 100

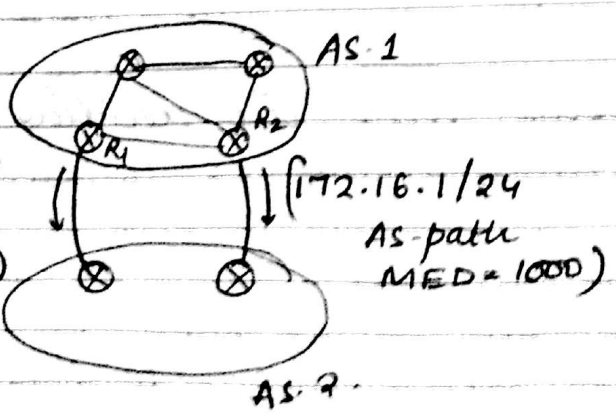
- R_2 tells BGP speakers (within AS 2)
  172.16.1/24 (AS 1, ...) local Pref = 60.

BGP Speakers : Routers Running BGP.

Note: <u>larger local-Preference</u> path is preferred.

* More autonomy ⇒ More freedom to run the AS administrator an its own.
* AS, have the flexibility to set the local preferences locally and later use them for comparisons & prioritizing.

2) MULTI-EXIT DISCRIMINATOR (MED)



- AS 1 hints AS 2 to use R2 to forward packets to this prefix 172.16.1/24.
- But AS 2 neednot follow. Still can send ~~on~~ to R1

Note: Smaller MED path is preferred.

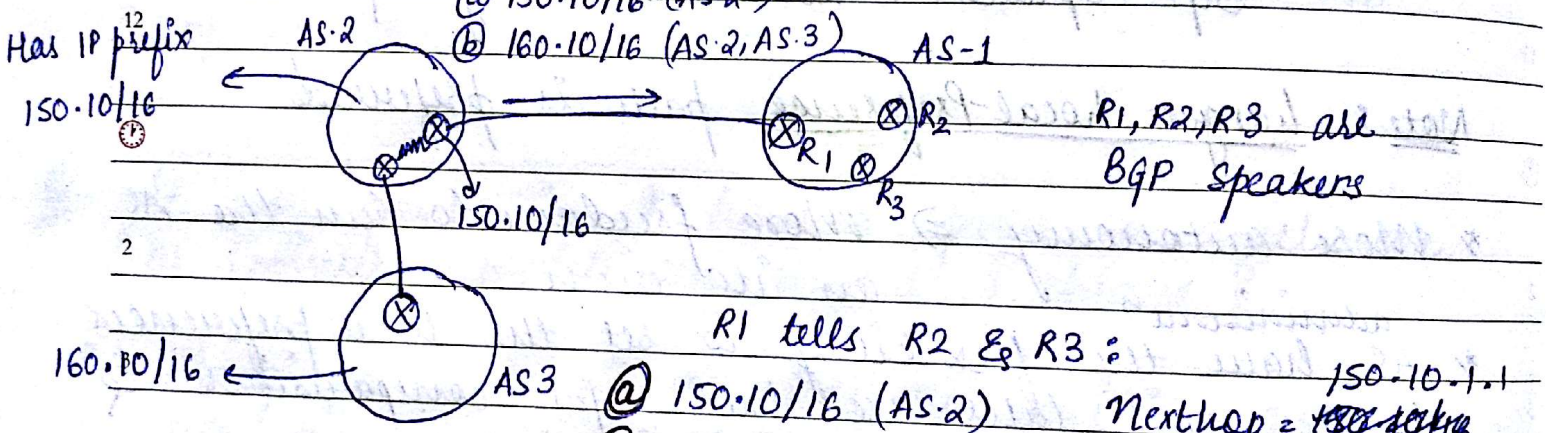
MESSAGES

3) AS-path

List of AS's on path to reach the advertised prefix.

4) NEXT HOP:

IP Address of external BGP router from whom the advertisement was received.



R1, R2, R3 are BGP speakers

R1 tells R2 & R3:

@ 150.10/16 (AS-2) Nexthop = 150.10.1.1

ⓑ 160.10/16 (AS-2, AS-3) Next hop = 150.10.1.1

↓

where should the pkt go next to reach the destinat".

Q How do these attributes work in deciding the preference?

⇒ Each BGP speaker decides routes as follows:

@ Choose largest local preference.

ⓑ Choose path with shortest AS-path.

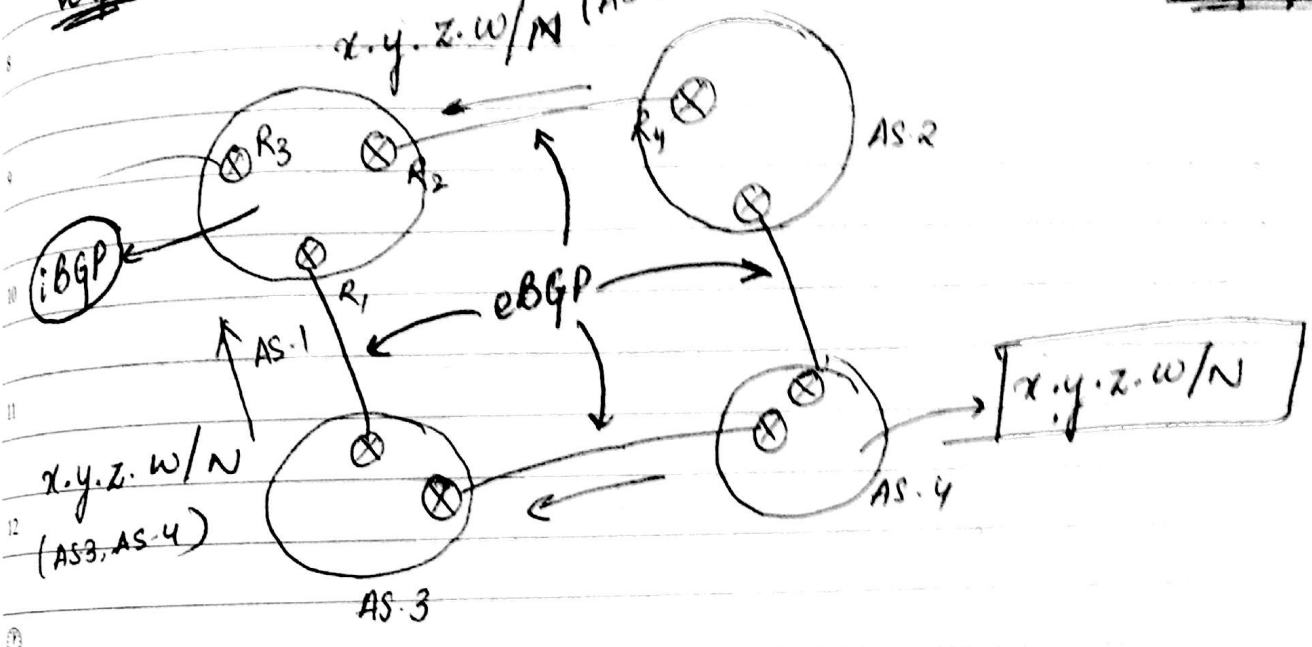
© Choose path with lowest MED. & to choose among paths from same neighbour (AS) &.

August'16

FRIDAY 17 JUNE

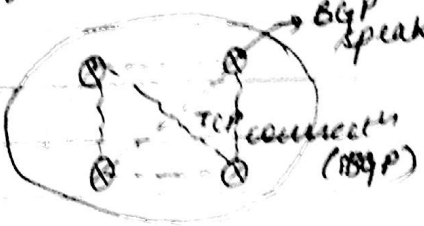
02/11/16

BGP:



→ Contd.

1) eBGP speakers learn AS-level paths from the neighbouring AS's.
2) eBGP speakers use iBGP to share routes with each other {normally full mesh of TCP connn's}
3) each BGP router selects routes for various IP prefixes.



eBGP, iBGP

IGP: Intrion Gateway Protocol (LSR / DV

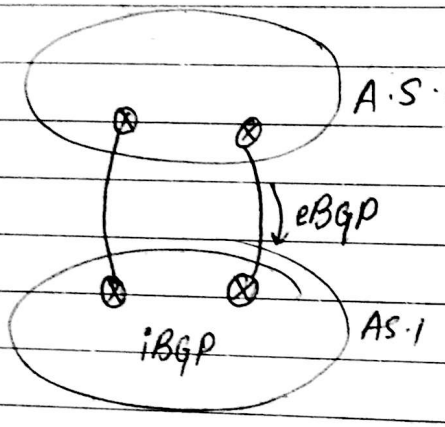
4) Insert chosen routes into IGP.
{So that non-BGP routers forward packets with external destinat" IP address correctly}.
5) eBGP speakers can advertise newly-created routes to other ASs

MESSAGES

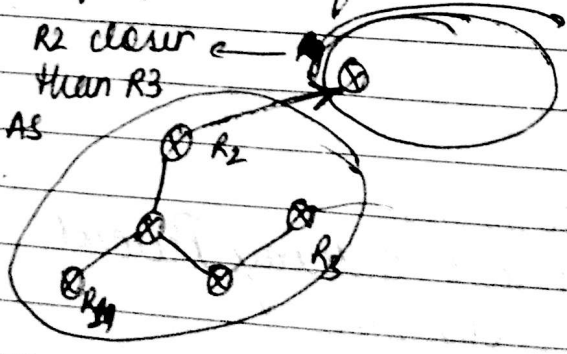
2016

Rules for choosing route { used by each BGP router }

1) Prefer Largest LOCAL-PREF.
2) Prefer shortest AS-PATH.
3) Prefer lowest MED.
4) Choose eBGP learned path over iBGP learned path.



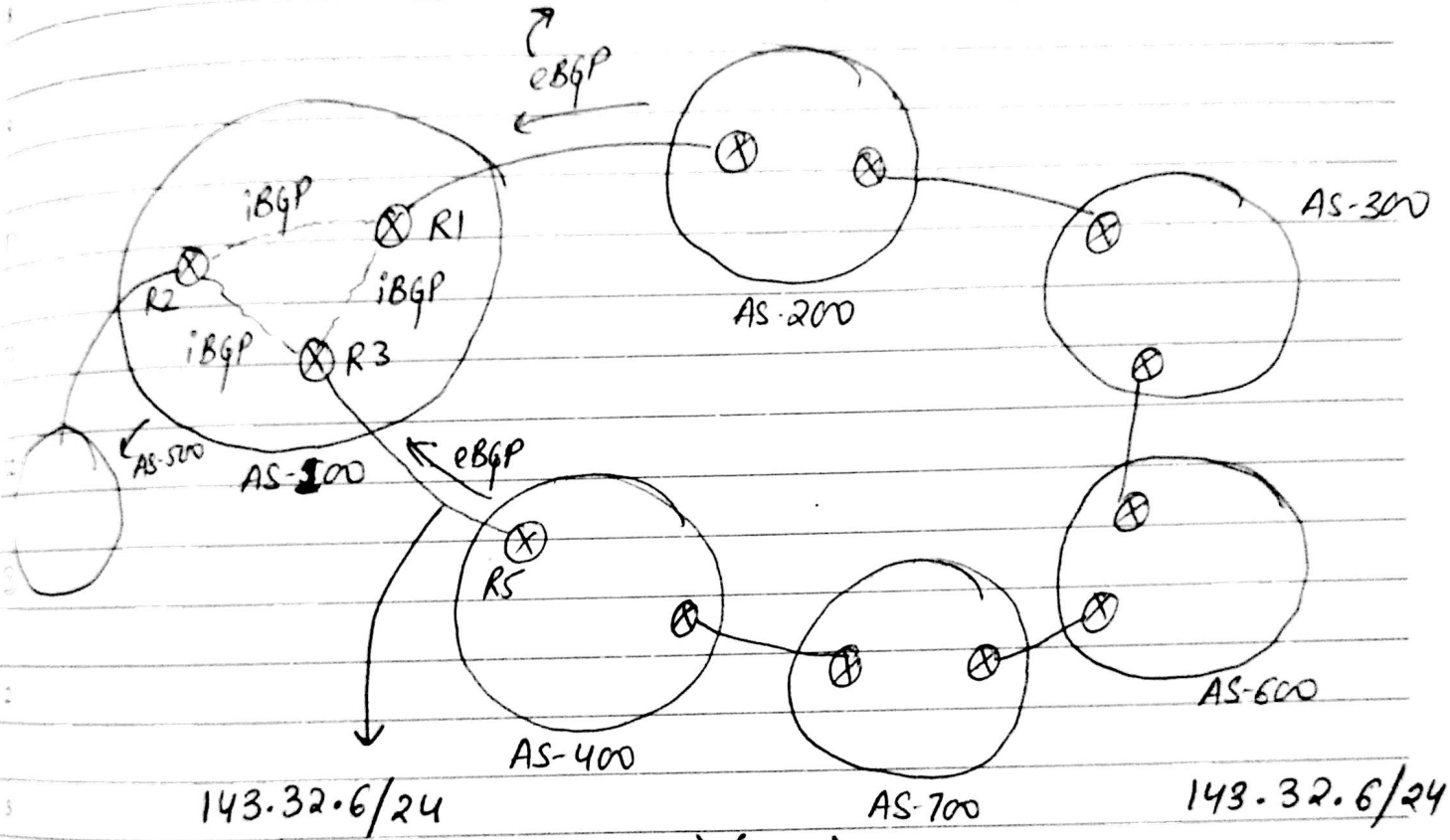
5) Hot Potato Routing
- Get rid of pkt. as fast as possible.



↳ Choose path with lowest IGP metric to Next hop

6) Router ID (lowest?)

Example: Advertisement of the prefix 143.32.6 (AS-200, AS-300, AS-600).



(AS-400, AS-700, AS-600) (... ...)

i) Suppose I want R1, R2, R3 to all use AS-200, AS-300, AS-600 ⇒ Set local-preference to make it largest

ii) Want AS-400, AS 600, for R1, R2, R3.

iii)

iv) eBGP also gives unique choice

R1 uses top path, R2 uses lower path.
R3 uses hot potato

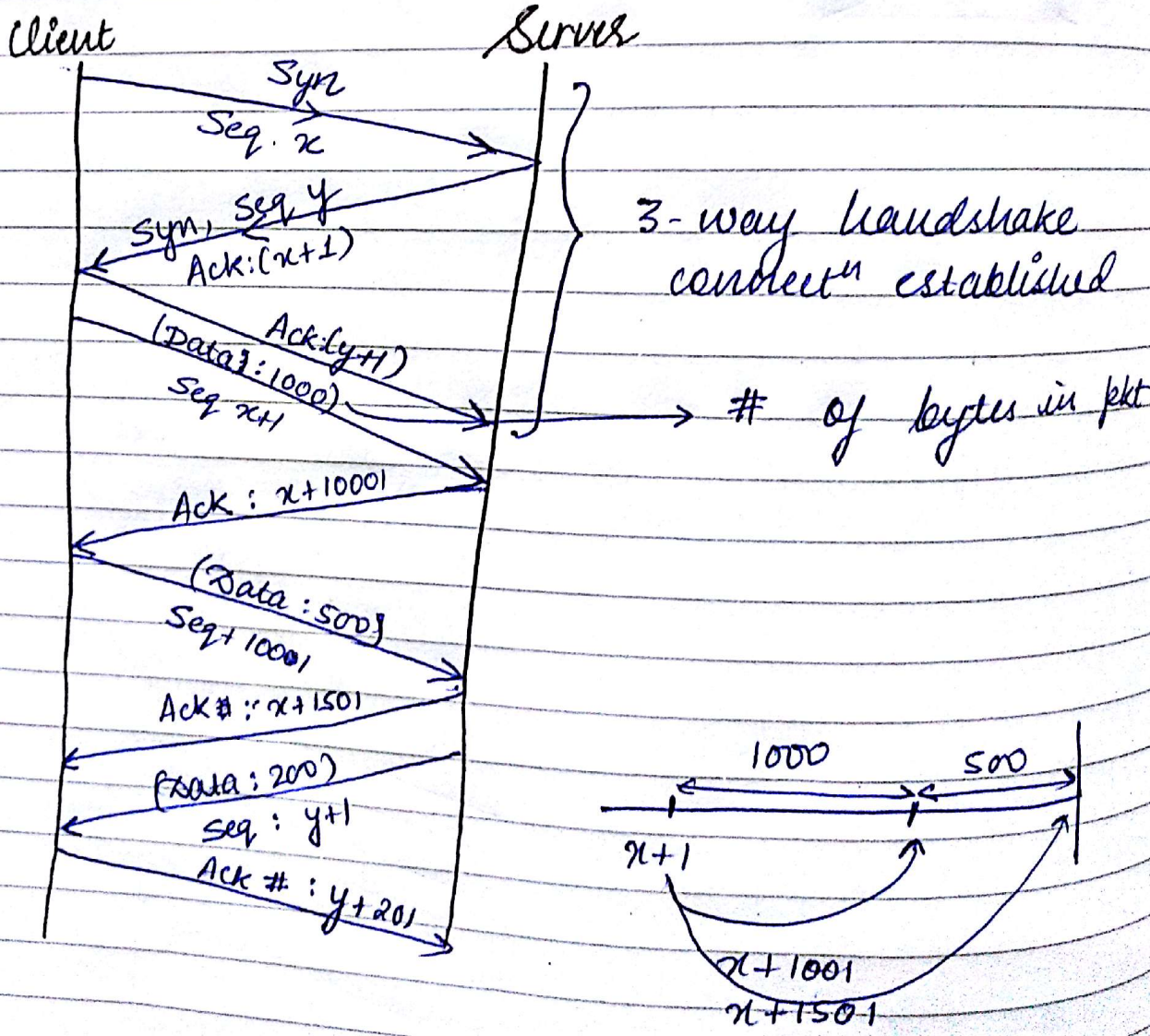
MESSAGES

2016

Transmission Control Protocol (TCP)

Appl.
↓
UDP TCP → Reliability
→ Congestion & Flow Control
→ In Order Delivery.

↓
Simple, Fwd-to Appl. based on port #



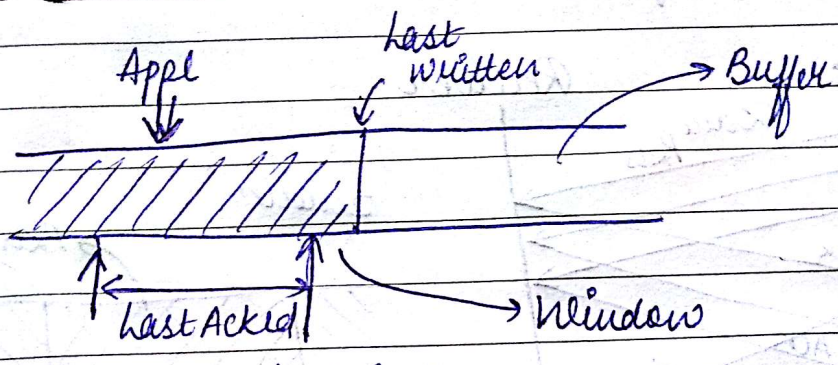
August'16

MONDAY 27
JUNE

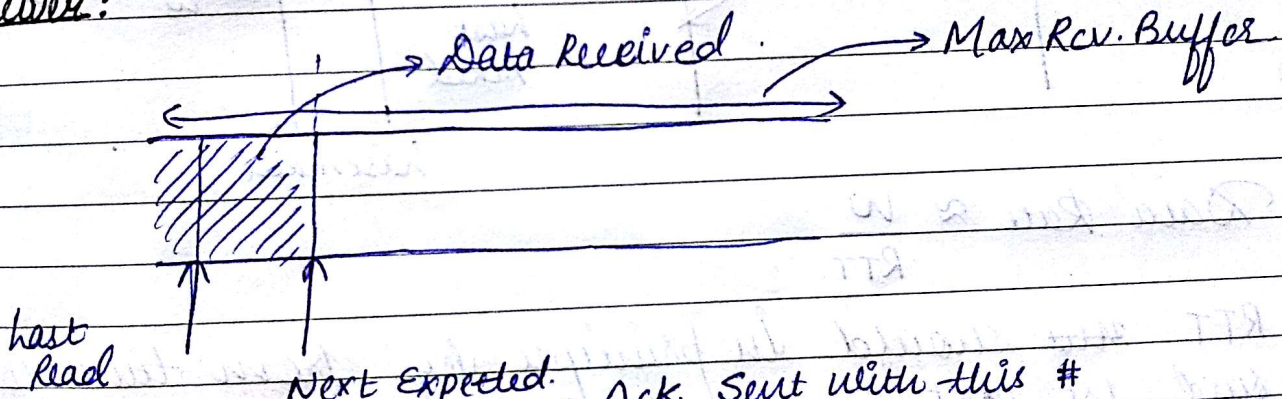
179-187 WK-26

Congestion Control:

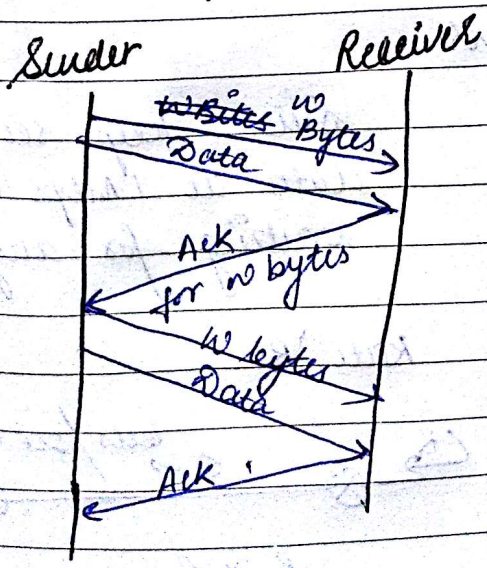
Sender:

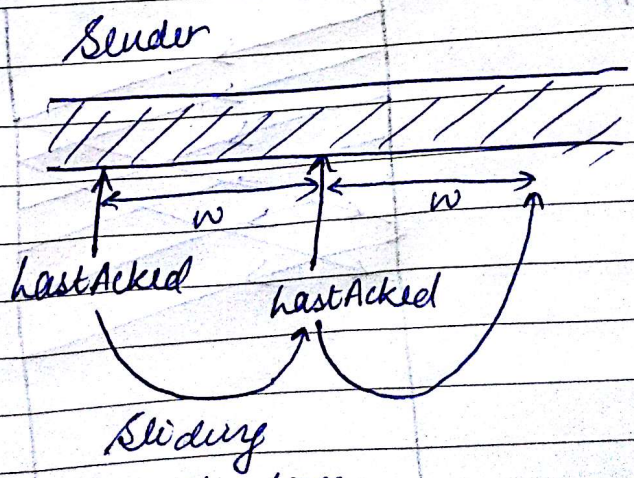


Receiver:



Significance of Window (W)





MESSAGES

2016

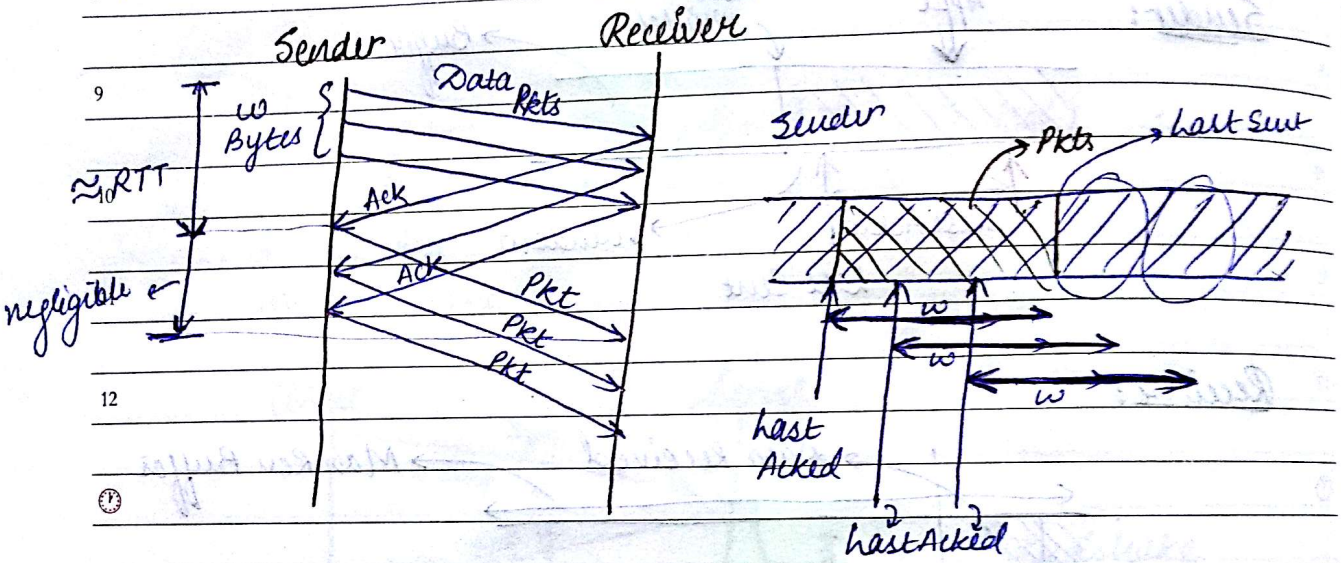
May'16

M	T	W	T	F	S	S
30	31					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

June'16

M	T	W	T	F	S	S
		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			

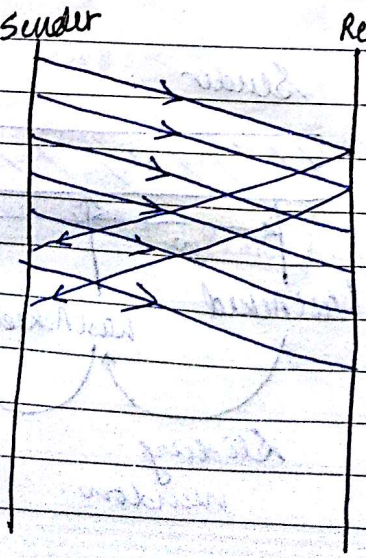
Assume 3 packets of data →



Data Rate ≈ $\frac{W}{RTT}$.

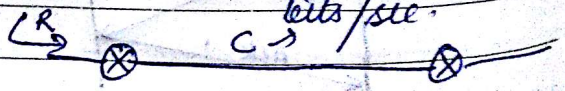
RTT should In principle be from time to send 1st pakt to the time 1st Ack is secived. It is almost same as till the time Receiver receives its pkt.

Suppose



Sender keeps sending data at 1 Mbps without waiting for any Ack.

Rate 'R'



$R \approx C$

$R = C + \varepsilon$

- Outstanding (Unacked) Data still in Network from Sender to Receiver.

$W = \min$ (Congestion Window, Advertised Window)
- Congestion Window → Control Congestⁿ in Network
- Advertised Window → Control Congestion at Receiver (Flow Control)

Flow Control

Advertised Window = [MaxRcvd Buffer − (NextExpct − LastRead − 1)]