Date _____

CHAPTER # 05

The Network Layer : Control Plane

→ The computation and maintainance of routing and forwarding table can be done in 2 ways :

① Per-router control :

→ Each router has :
- A routing function (runs routing algo like OSPF, ~~BGP~~)
- A forwarding function

→ Routers communicat with each other to share info and independently compute their own forwarding tables.

② Logically Centralized (SDN) ~~Approa~~ Control

→ A central controller computes all routers's forward table

→ Router become simple forwarding device that follow instructions from controller

→ Each router has a CA (control agent) that interacts with controller thru some protocol.

Date _____ Day _____

▷ Routing Algorithms

↳ goal is to calculate the least cost path from a sender → receiver

- Types Of Routing Algorithms

① Centralized vs Decentralized

→ Centralized:

- Global knowledge of all nodes and link costs
- Calculates least-cost path using complete network topology
- Can be run at a central controller or at all routers
- Ex: Link-state algo

→ Decentralized:

- No global knowledge; each router knows only its direct neighbour
- Routers exchange info iteratively to compute paths
- Each router maintains a distance vector (cost to all nodes)
- More suited to distributed control

Date _____ Day _____

(2) Static vs Dynamic

→ Static

- Routes change rarely, usually via manual changes
- Not adaptive to real-time conditions

→ Dynamic Routing

- Routes automatically change based on topology or traffic
- Can be periodic or event-triggered
- More responsive, but can face loops and oscillations

(3) Load Sensitive Vs Load Insensitive

→ Load Sensitive

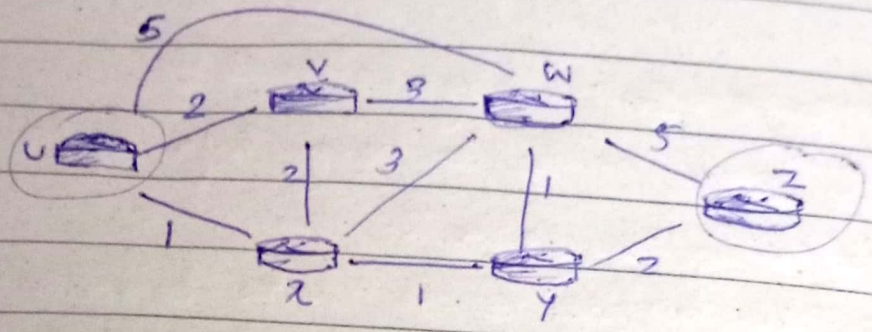
- Link cost vary with traffic congestion
- Tries to avoid congested links

→ Load Insensitive

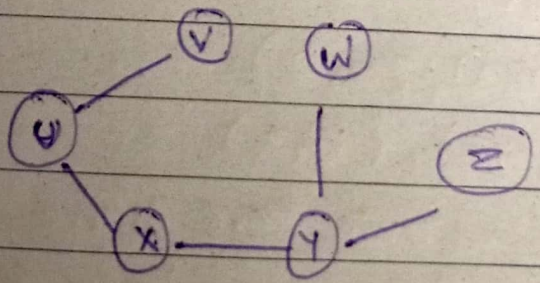
- Link cost is fixed, doesn't reflect congestion
- Used in most current internet protocols like RIP, OSPF, BGP

Date _____ Day _____

- Link-State Algorithm
  - ↳ It is a dijkstra algo



	v	x	y	w	z
u	2,u	1,u	∞	5,u	∞
ux	2,u		2,x	4,x	∞
uxv			2,x	4,x	∞
uxvy				3,y	4,y
uxvyw					4,y
uxvywz					



• go v par jana hai u se to u,v hi hop lena parega

→ Forwarding Table

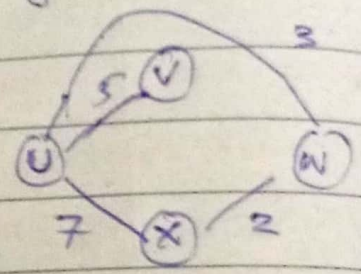
Destination	Link
v	u,v
x	u,x
y	u,x
w	u,x
z	u,x

• in mein se kahi bhi jana hai to next hop u se u,x hi hoga.

Date

- Complexity:
  → $n(n-1)/2$ comparision → $O(n^2)$

- Message Complexity:
  → LS algo needs to have info of all the nodes in the network.
  → For this purpose each router broadcast it's link state information to other 'n' routers connected to it.
  → In reports se network topology ban sakti hai.
    ↳ Node U sending LS report
      Node U: (v,5), (w,3), (x,7)
    ↳ Node U received following LS reports
      Node V: (U,5)
      Node W: (U,3), (x,2)
      Node X: (U,7), (W,2)



  → Therefore by this all node have entire topology info.

route
- Oscillation possible when link cost depends on traffic volume
  ↳ prevention: All routers donot run LS algo at same time, all routers donot broadcast LS report at same time

Date _____ Day _____

- Distance Vector Algorithm
  - ↳ It is based on Bellman - Ford algo
  - ↳ DV algo is iterative, distributive and async
    - Process continues until no more info is being exchanged
    - Each node receives info from neighbor, calculates path, send info back to neighbor
    - doesn't req all routers to perform calc sequentially!

Bellman Ford eqn

direct cost from $x \to v$

$$D_x(y) = \min_v \{ c_{x,v} + D_v(y) \}$$

- $D_x(y)$: cost of least cost path from x to y
- $c_{x,v}$: x se uske neighbour ~~ka~~ jiska neighbour se min cost hogi wo lenge
- $D_v(y)$: neighbour se destination ki cost

$$D_u(z) = \min \{ c_{u,v} + D_v(z), c_{u,x} + D_x(z), c_{u,w} + D_w(z) \}$$
$$= \min \{ 2+5, 1+3, 5+3 \}$$
$$= 4$$

Each Node:

- wait for change in local link cost or an updated DV from neighbor
- recompute DV estimates using DV recv from neighbor

Page No. ☐

Date _____ Day _____

- if DV to any destination change only then notify neighbors

→ See example from slides at 36...

- Effect of Link cost changes:
  - ↳ Node detects local link cost change
  - ↳ Updates routing info, recalculates DV
  - ↳ if DV changes notify neighbors

→ poisoned reversed (video)

Comparasion of LS & DV

① Message Complexity
- LS : with n nodes, E links O(nE)
- DV : Exchange b/w neighbors only

② Speed of Convergence
- LS : $O(n^2)$ algorithm, may have oscillations
- DV : convergence time vary ; agr congestion hoa to routing loops or count to infinity problem hosakti hai.

Da → What happens when router malfunctions Day

③ Robustness

- LS:
  - node can advertise incorrect link cost
  - each node computes its own table indep.

- DV:
  - DV node can advertise incorrect path cost
  - each node's table is used by other
    ↳ error propagate thru network

- Poisoned Reversed:

→ If z routes through y to get to destination x, then z will advertise to y that its distance to x is ∞, i.e z will advertise to y that $D_z(x) = \infty$.

Now y believes that z has no path to x, y will never attempt to route x via z.

Date Day

D INTRA-AS ROUTING IN INTERNET : OSPF

→ Routers are grouped into regions known as Autonomous Systems a.k.a domain
→ Each AS is identified by a unique AS no. (ASN)
→ Routers in an AS can use same routing protocol
→ Most of the time each ISP constitutes a single AS or there might be multiple AS in a single ISP

- intra-AS (intra-domain)

→ Routing among routers within same AS. ("network")
↳ All router in AS must run same intra-domain protocol
↳ routers in diff AS can run diff intra-domain routing protocols
↳ gateway router : at edge of its own AS, has links to routers in other AS'es

- inter-AS (inter-domain)

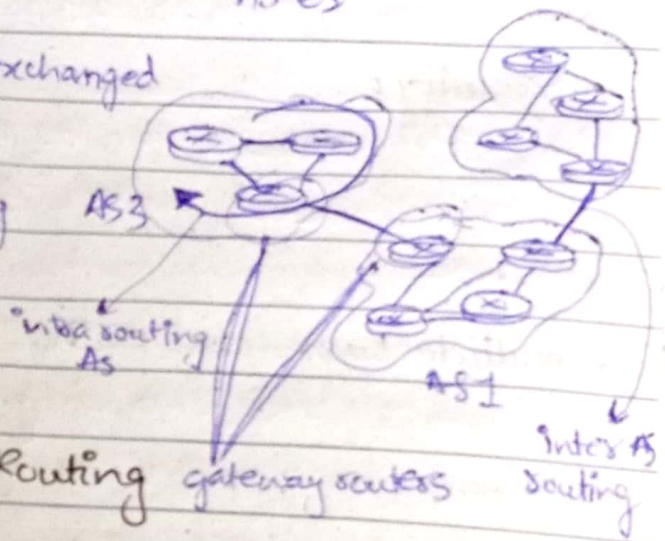
→ Routing amongst AS'es
↳ gateway perform inter-domain routing (as well as intra-domain routing)

Date _____ Day _____

- Most Common Intra-AS Routing protocols

① RIP : Routing Information Protocol
→ Classic DV : DV exchanged every 30 second
→ no longer widely used

② EIGRP : Enhanced Interior Gateway Routing Protocol
→ DV based



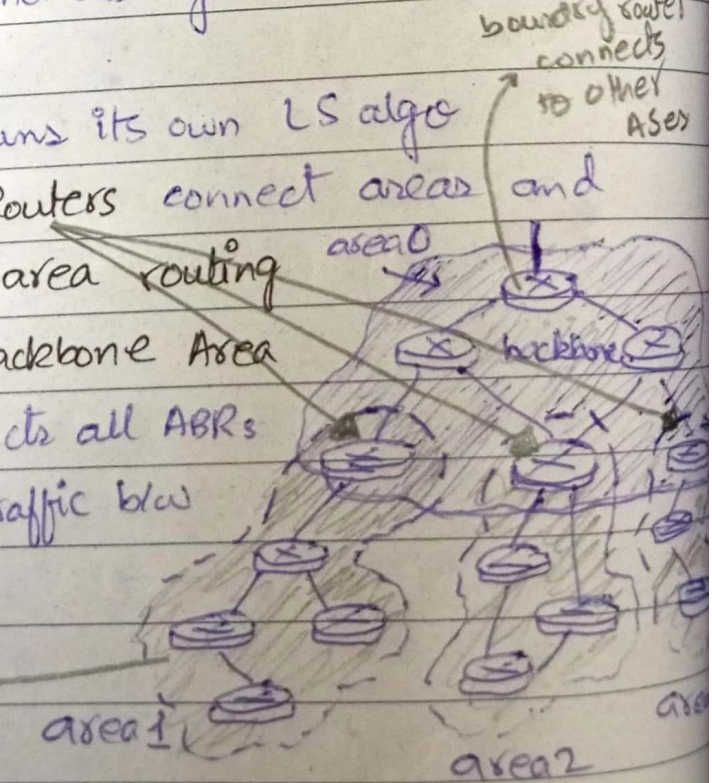
③ OSPF : Open Shortest Path First
→ Classic link-state.
→ Each router floods OSPF link state advertisements (directly over IP rather than using TCP/UDP) to all other routers in entire AS.
↳ Therefore need to implement rdt itself
→ Each router has complete topology of the AS
→ Each 4 runs Dijkstra to determine shortest path to all routers/subnets
→ Link cost is set by network Admin, if it is set to 1 then this achieving minimum hop routing

with an upper layer protocol for OSPF

Date Day

→ Router broadcast Link-state info whenever there is change or periodically ~~after~~ at least once in 30 min.
↳ adds robustness

- Security:
  → OSPF packets can be authenticated to prevent unauthorized routers from injecting false routing info.
- multiple same-cost paths
  → Supports load balancing by allowing multiple equal-cost paths to a destination, instead of choosing just one.
- Hierarchical Routing (Intra AS)
  → An OSPF Autonomous System (AS) is divided into areas
  ↳ Each area runs its own LS algo
  ↳ Area Border Routers connect areas and handle inter-area routing
  ↳ A special Backbone Area (Area 0) connects all ABRs and routes traffic b/w areas.



Day

routers:
Flood LS in an area only
→ Compute routing within an area
↳ forward pkt to outside via area border router

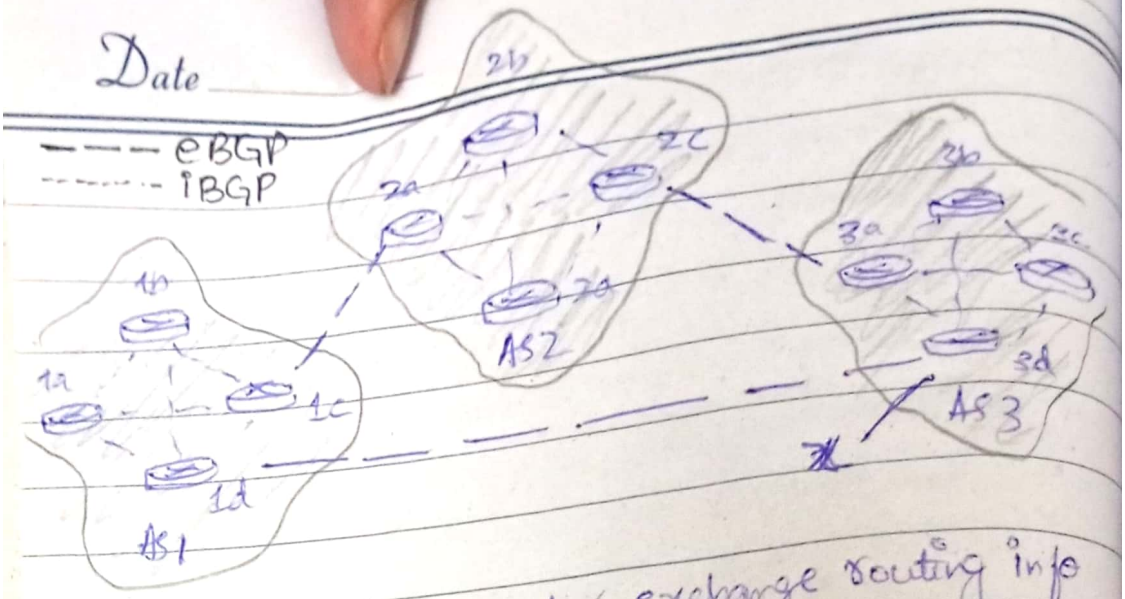
Backbone routers:
↳ runs OSPF limited to backbone

▷ Border Gateway Protocol (BGP)

→ BGP is used as inter-autonomous system routing protocol
→ Allow subnet to advertise its existence and the destination it can reach.
→ In BGP, pkts are not routed to a specific dest addr, but instead to a CIDRized prefixes, with each prefix representing a subnet or a collection of subnets, like 138.16.68/22
→ BGP helps to determine "best" routes to the prefixes

• Advertising BGP Route Information
edge of an AS directly connects to routers in another ASes
→ Each router in an AS is either a gateway router or an internal router
↳ connects only to hosts & routers within its own AS

Date Day



→ An BGP pairs of router exchange routing info over semi-permanent TCP connections port 179

→ This connection is called BGP connection

→ A connection b/w two gateway router is called external BGP (eBGP)

→ between internal routers → internal BGP (iBGP)

→ iBGP connection donot always corresponds to physical links

→ Sending reachibility info for prefix to all other routers in AS1 & AS2

① eBGP (Inter-AS)
- Router 3a send "AS3 x" to Router 2c

② iBGP (Intra-AS)
- Router 2c sends "AS3 x" to all routers in AS2

③ eBGP (Inter-AS)
- Router 2a send "AS2 AS3 x" to Router 1c

④ iBGP (Intra-AS)
- Router 1c sends "AS2 AS3 x" to all routers in AS1

Day

- Determining Best routes

→ "Route" in BGP
- A route = Prefix + Attributes
- Important attributes
  - → AS-PATH
  - → NEXT-HOP

→ AS-PATH Attribute
- List the sequence of AS no.s a route has passed thru.
- Example:
  - For prefix "x", AS1 might see
    - AS2 AS3 x (via AS2)
    - AS3 x (via AS3)
- Used for path selection and loop prevention (if a routes sees its own AS no. ,it rejects the route)

→ NEXT-HOP Attribute
- Indicates the IP Addr of the next-hop router interface that starts the AS-PATH.
- Example:
  - For route "AS2 AS3 x", NEXT-HOP = IP of left interface of 2a

Date Day

- For route "AS3 x", NEXT-HOP = IP of left int of router 3d
- NEXT-HOP is the IP of router that does not belong to AS1.

→ Hot Potatoe Routing
→ If we want to reach 1b → x that will 2a, 3d
→ We check all possible gateways of AS1.
→ Using routing info from intra-AS protocol to determine cost of least-cost path to each of gateways.
↳ For 2a it will be 2 hops
↳ " 3d " " " 3 hops
→ Hot potatoe routing: Choose the gateway that has the smallest least cost
→ Determine the forwarding table the interface I that leads to 2a.

→ BGP Route Selection
→ If there are more than one route to a destination then BGP applies below rules until only one route remains.
① Each route has a local prefrence attribute set by Network Admin. The route with highest val is selected

Date _____ Day _____

② If all have same local preference value, then one with shortest AS-PATH is selected

③ If all have same local pref. & AS-PATH length, then hot-potatoe is used

④ If still more than one route still remains, the router uses BGP identifiers to select the route.

Why Different intra-, inter-AS routing?

- Policy:
  - inter-AS: admin wants control over how its traffic routed, who routes thru its network
  - intra-AS: single admin, so policy less of an issue
- Scale:
  - hierarchical routing saves table size, reduced update traffic
- Performance:
  - intra-AS: can focus on performance
  - inter-AS: policy dominates over ↓

Date _____ Day _____

CHAPTER # 06

— The Link Layer and LANs —

- **Node:** Any device running a link-layer protocol
  eg (host, routers, switches, WiFi access point
- **Link:** The physical communication channel connecting adjacent nodes

→ A datagram travel from src → dest accross multiple links

→ Over each link, the datagram is encapsulated in a link layer frame and transmitted

▷ Services Offered by Link Layer

- Flow control: pacing blw adjacent sending & recv nodes

① **Framing**

- Encapsulates the network-layer datagram into a link-layer frame

② **Link Access (MAC - Medium Access Control)**

- MAC is a protocol that specifies the rule by which a frame is transmitted onto the link.

③ **Reliable Delivery**

- Ensure error free transmission over the link using ACKs & retransmission

Date _____ Day _____

- Common in wireless links with high error rates
- Often skipped in wired links due to low error rates

(4) Error Detection and Correction

- Detect bit errors using error-detection bits (like checksums)
- Error correction not only detects but also corrects the errors
- Typically implemented in hardware for efficiency

▷ Implementation of Link Layer

- Mostly in Hardware
  → Implemented in a network adapter (NIC - Network Interface Controller)
  → Integerated into motherboard or on a seperate chip
  → Handles core function: Framing, link access, error detection
- Partly in Software
  → Runs on Host CPU
  → Responsible for
    ↳ Managing addressing info
    ↳ Activating controller hardware
    ↳ Responding to Interrupts
    ↳ Passing recv data to the network layer

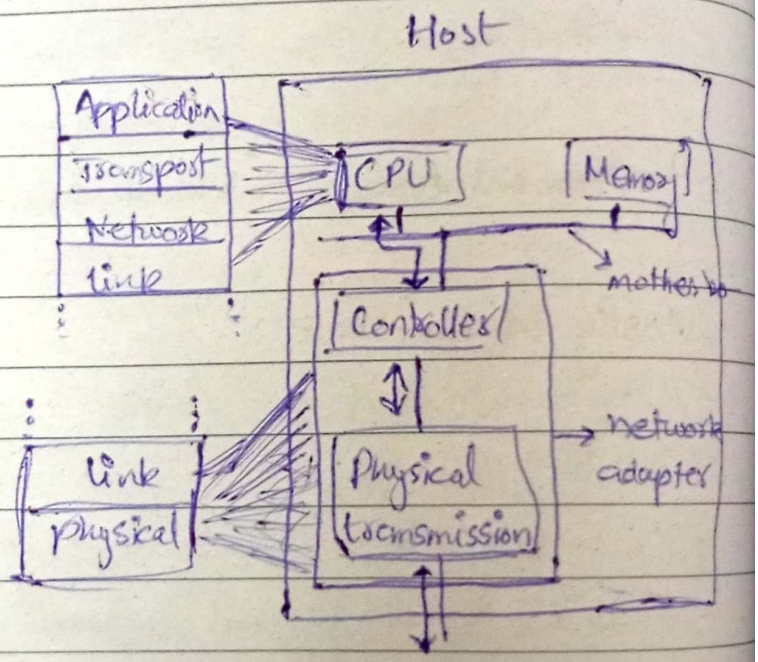
Date _____ Day _____

- Send side workflow:

  Host CPU builds a datagram → NIC encapsulates it in a frame → Sends over the link

- Recv side workflow:

  NIC recv frame → check for errs → extract datagram → hands it to software → software sends it to network layer.



▲ Network adapter

Day

- MAC Addresses
  - → Adaptors (NIC) of hosts & routers have MAC Addresses
  - → Switches do not have MAC Addr on their interfaces because they forward frames transparently
    - Eg, k host or routers explicitly switch k interface ko address nhi kate to send datagram.
  - → 6 bytes (48 bits) long
  - → Expressed in hexadecimal (1A-2B-3C-4D-5E-6F)
  - → Assigned by manufacturers & managed by IEEE
  - → Unique & flat (non-hierarchical) structure
- MAC Vs IP
  - → MAC : Fixed, unique/adapter (NIC)
  - → IP : Changes with location / network (postal addr)

- Sending Frames
  - → Sender includes destination MAC addr in the frame
  - → The frame is recv by switch which occasionally sends this frame to all connected devices
  - → The recving adapter
    - ↳ Accepts the frame if dest MAC matches from it's own
    - ↳ Discard otherwise

Page No.

Day

- Broadcast Address
  - → Special MAC addr : FF-FF-FF-FF-FF-FF
  - → Used when a frame must be delivered to all devices on the LAN.
  - → Common in protocols like ARP or DHCP.

▷ Address Resolution Protocol

→ ARP is used to map an IP Addr to MAC Addr within the same subnet

→ Essential for hosts/routers to send Link-layer frames when they know the IP but ~~not the~~ not the MAC addr of the dest

- How ARP Works
  - ARP Table (cache)
    - → Each host/router maintains an ARP table IP → MAC mapping with TTL (time to live)
    - → Not all mappings are pre-populated ; entries expire (typically after 20min)

- When Sending to a Known IP (and MAC is in ARP table)
  - → Sender looks up MAC of dest IP in ARP table
  - → Encapsulates IP datagram in a frame with dest MAC
  - → Sends it directly over the link

broadcast at link layer
to all recv

Day

- When MAC is unknown
  - → Sender creates an ARP request (query) pkt
    - ↳ contains sender's IP & MAC and target IP (whose MAC is needed)
  - → Sent in a broadcast frame with MAC dest FF-FF-FF-FF-FF-FF
  - → All nodes on the LAN recv the broadcast frame
  - → Host with matching IP
    - ↳ Recognizes the match
    - ↳ Replies with an ARP reply containing
      - its MAC Addr & IP
    - ↳ Reply is sent in a unicast frame (not broadcasted)

Sent as unicast (directly to the requester's MAC)

  - → Sender recv ARP reply
    - ↳ Updates ARP table with the new IP-MAC mapping
    - ↳ Uses the MAC to send datagram.

Q: Why IP & MAC Add needed when there is IP?

→ In a LAN a switch doesnot recognize the IP it only recognize the MAC.

→ IP are used to route pkts end-to-end accross network

→ MAC are used to deliver frames within LAN

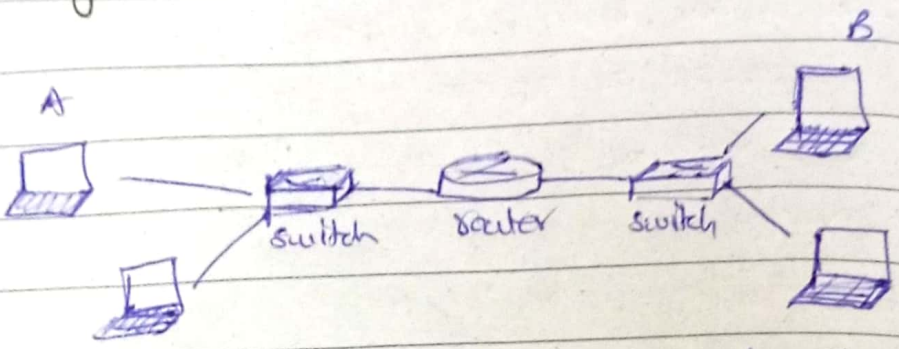
Page No.

Day

Date

because the data will actually be recv by NIC, which only understands MAC.

▷ Routing to another Subnet



① "A" creates IP datagram with IP src A, dest B
② "A" creates a link layer frame containing A to B IP datagram
③ For dest MAC, it will be R's left interface MAC, not B's MAC. Because A doesn't know B's MAC.
④ Frame sent from A to R
⑤ Frame recv at R, datagram removed, passed upto IP
⑥ Router determines outgoing interface through it forwarding table by seeing dest IP.
⑦ Router then creates link layer frame containing A to B IP datagram.
⑧ Frame dest MAC will now be B's MAC (router will find it thru ARP).
⑨ Transmits link layer frame

Date

(iii) "B" recv frames, extracts IP datagram

(iv) "B" passes datagram up protocol stack to IP