

CS348 Notes  
BGP+IGP  
Video Numbers: 20, 21

OjMaha

I have prepared these notes by watching the videos from [Networks Playlist](#). The following notes may be asynchronous and irrelevant to what Prof. Vinay teaches in class (cuz I do not pay attention during lectures lol). Further, these notes might not cover *everything* as explained in the video lectures. Consider these to be a supplemental read :). If you find any errors, do notify me so they can be edited.

# Border Gateway Protocol (BGP)

It conducts inter-domain routing b/w ASes.

**Customer:** The one who pays for service. (Multi-home customer: pays multiple providers for service)

**Provider:** The one who provides service.

**Tier-1 AS:** pays nobody. It only acts as a provider. They are widely connected

**Tier-2 AS:** takes internet service from Tier-1 AS.

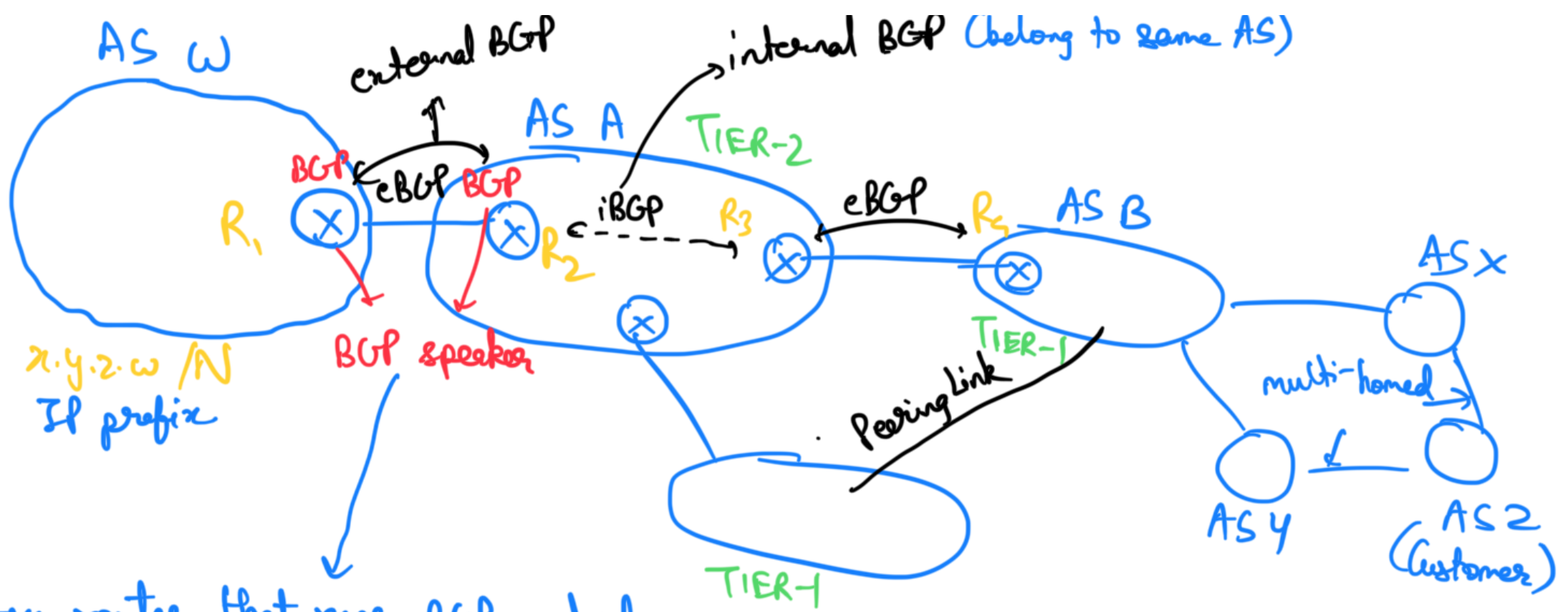
Define Tier-n likewise.

**Peering Link:** Connection b/w two Tier-1 AS for convenient data sharing to others.

No one pays the other here. Since this is unpaid, the bandwidth is low to avoid "cheating" by sending huge amts of data to the other Tier-1. It is "under-provisional".

**Service Level Agreement:** A QoS guarantee b/w 2 AS abt latency, bandwidth.

Note that there cannot be a full end-to-end latency<sup>/bandwidth</sup> guarantee since AS do not know the whole path. The agreement is just for within own network.



Any router that runs BGP protocol.

$R_1$  tells  $R_2$  :  $x.y.z.w/N$  (IP prefix)  $W$  (AS Path)  $\underbrace{\text{attributes}}$   $\rightarrow$  Advertisement.

$R_3$  tells  $R_4$  :  $x.y.z.w/N$  (IP prefix)  $A-W$  (AS Path)  $\underbrace{\text{attributed}}$

$R_5$  tells  $R_6$  : — —  $B-A-W$  (AS Path)

The AS path is the path from me to reach the IP prefix. It also means that if  $R_6$  gives  $R_5$  an IP pkt with Dest. IP matching prefix;

it will be able to route it along B-A-W.

⚡ B does not necessarily have to tell C that there is a path B-A or B-A-W.  
B & C can decide to keep some of their connections quiet. Because both are Tier-1 AS, they do not want the other to send its traffic to it (B).  
(say B) (say C)

Similarly, Z is a small network and doesn't want packets from X to Y to be routed through it. It doesn't advertise the connection.

BGP does not compel any router to share its ads.

e-BGP: protocol b/w speakers of diff. ASes. They use TCP port 179. no. of routers in AS.  
i-BGP: same AS.  $\nexists$  pairwise TCP connections.  $O(n^2)$  connections.

IGP: Interior Gateway Protocol (Intra-domain routing)  $\Rightarrow$  LSR, DV, etc.

Steps to decide e-BGP paths:

- (i) e-BGP speakers learn AS-paths from neighboring routers in other ASes.
- (ii) - and also learn learned info via iBGP with other BGP speakers in own AS

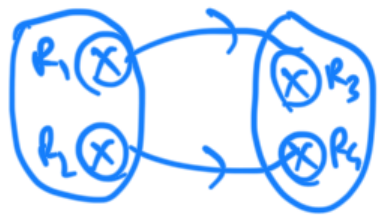
- (ii) e-BGP makes some routes...
- (iii) BGP speakers select routes to various IP prefixes.
- (iv) Insert chosen routes into IGP. We do this cuz there may exist nodes & speakers that do not follow BGP but only IGP. Thus this allows all routers to fwd packets to a destination IP.
- (v) eBGP speakers can now advertise newly created routes to neighboring ASes.

### BGP attributes:

#### (i) Local-preferences

Suppose AS2 receives two ads for same prefix. The network administrator determines what the local pref for these routes. Higher local pref  $\Rightarrow$  more preferred path.  
This pref. gets forwarded.

#### (ii) Multi-exit discriminator (MED)



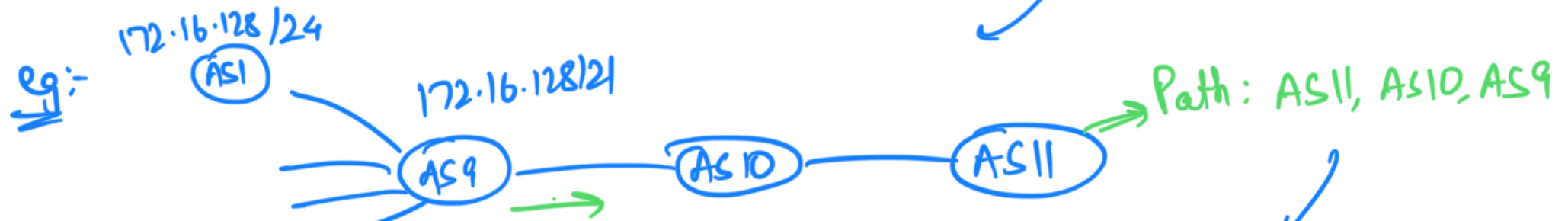
Suppose  $R_1 \rightarrow R_3$  &  $R_2 \rightarrow R_4$  advertise abt. same prefix but diff. MEDs.  
 $\downarrow$  MED: 2000       $\downarrow$  MED: 1000  
 say P

This means; AS1 is telling AS2 that it prefers packets with dest. P to be routed via  $R_4 \rightarrow R_2$ . Lower MED  $\Rightarrow$  higher pref. The value may be the IGP distance



to the next BGP speaker.

(iii) AS Path: List of all AS nos all the way to the dest. with that IP prefix.  
 whoever first advertises that prefix.

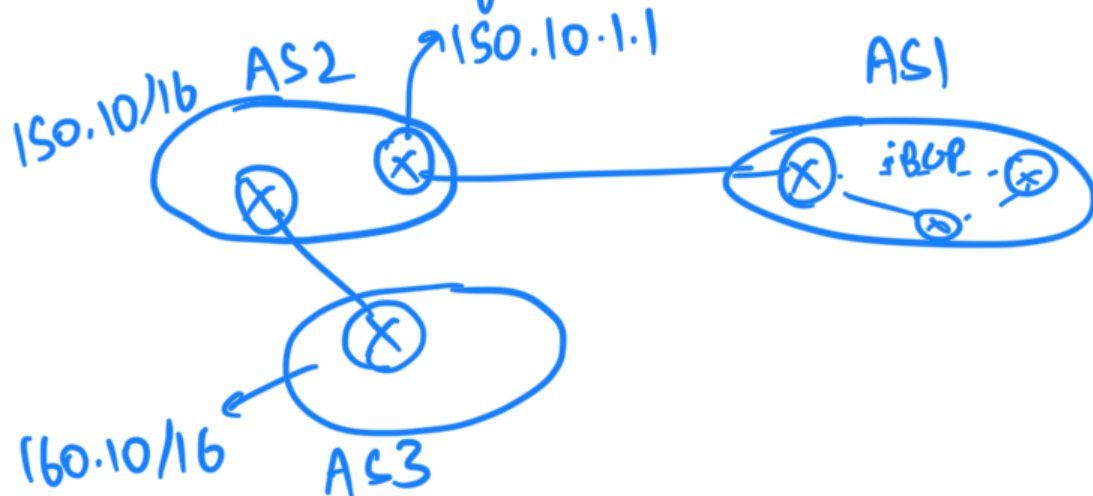


if the dest. is AS1 but AS9 is the first who advertised, the path will not contain AS1 but end at AS9.

After AS9 receives info; it forwards to AS1.

(iv) Next-Hop:

IP address of the first router in the next AS.



info that AS1 has: Next-Hop

Prefix	AS	Next-Hop
150.10/16	AS2	150.10.1.1
160.10/16	AS2, AS3	150.10.1.1

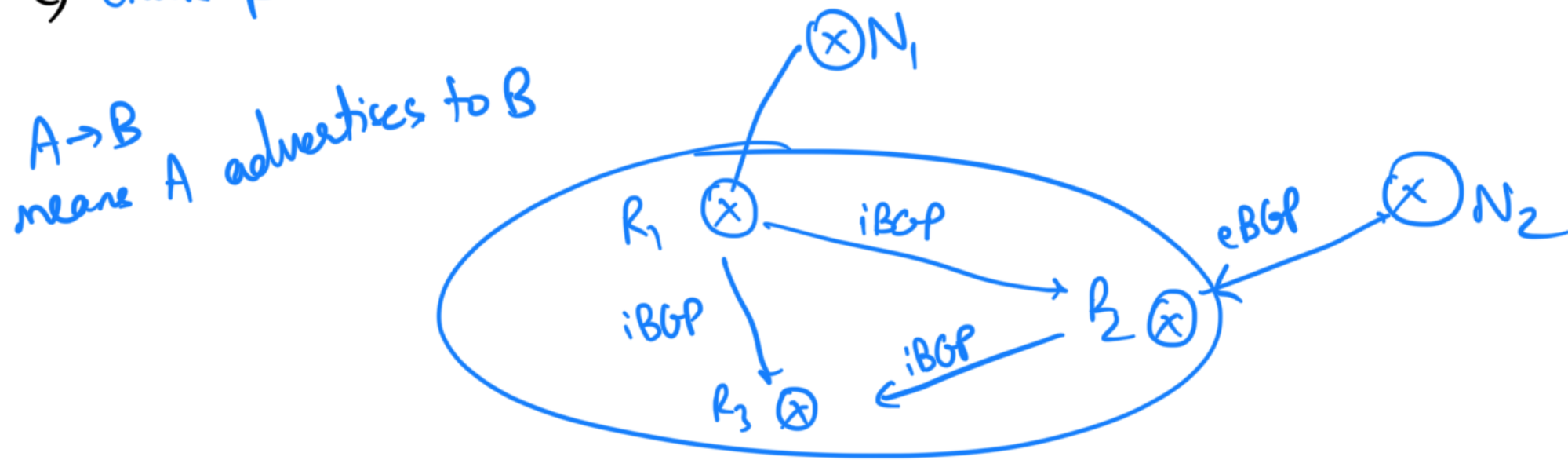
dest? Path

Rules to choose Routes:

AC route to use among those available

Each BGP speaker must decide which AS-path to use among many alternatives for the same prefix. (tie-breaks along the way)   
 if no local pref was set   
 tie-breaker for  $a_1 = a_2$  &  $b_1 = b_2$

- Use largest local-pref.
- Use shortest AS-path.
- Choose lowest MED path
- Choose path learned by eBGP over path learned by iBGP.
- Choose path with lowest IGP metric to next-hop.  $\rightarrow$  HOT-Potato routing



Here,  $R_2$  chooses to route over  $N_2$  rather than  $R_1$  in case # AS, and all other metrics are the same.   
 (eBGP) (iBGP)

Here,  $R_3$  recd only iBGP from  $R_1$  &  $R_2$ . Say  $\text{dist.}(R_3, N_2) > \text{dist.}(R_3, N_1)$ .

Then;  $R_3$  chooses to route via  $R_2$ . (nearest exit) (garam also routing)

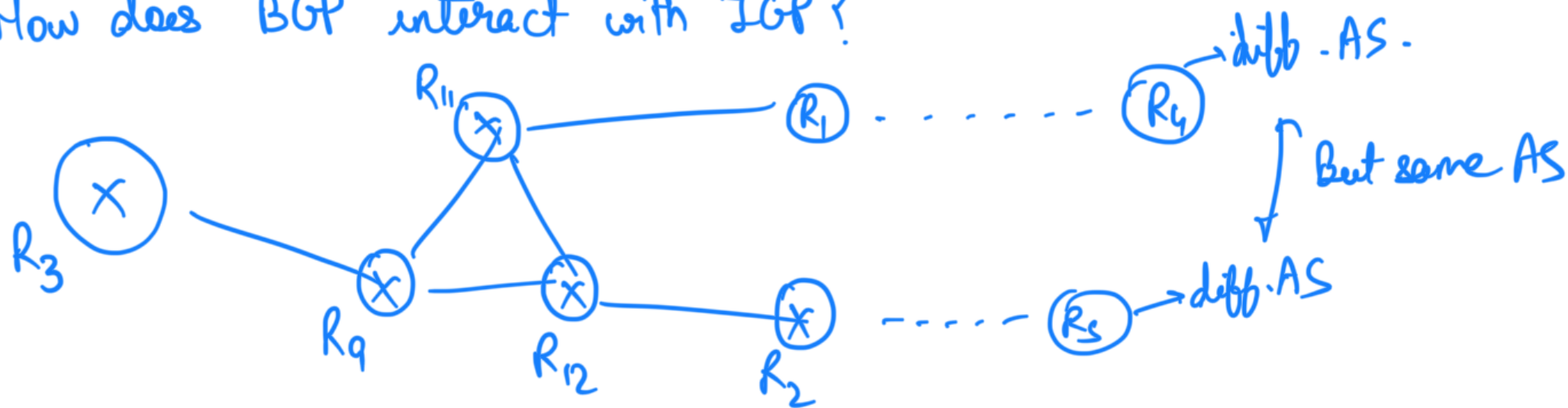
b) Choose lowest router ID among all BGP speakers who have sent their ads.

↳ router ID: highest IP on router. This is the final tie-breaker.

Why do routers have diff. IP addresses? Because the router may be a part of several networks, each of which requiring a diff. prefix. To satisfy all the cond<sup>n</sup> for those nets, it has diff. IPs allowing flexibility.

Watch vid lec 21 till 22 mins for complete example.

How does BGP interact with IGP?



(i) Encapsulation:

Suppose  $R_9, R_{11}, R_{12}$  have no BGP info. It only has info on how



to route pkt to routers of own AS.

in  $R_3$  table (IGP)

Dest.	Next Hop
$R_1$	$R_{11}$
$R_2$	$R_{12}$
$R_3$	$R_3$
$R_{11}$	$R_{11}$
$R_{12}$	$R_{12}$

Sidenote: each interface in a router has diff. IP.  
(Here, we ignore that)

Suppose dest is some router D with IP 142.32.6.21. It is not in routing table of  $R_3$ . When  $R_3$  receives this IP packet, it creates another IP pkt where the src packet becomes the payload.



$R_1$  receives this pkt, strips off the outer IP layer and forwards the internal pkt to  $R_4$ . ( $R_1$ 's table has info of  $R_4$  via eBGP)

## (ii) Pervasive BGP:

All routers must speak BGP.

Suppose there is a unique exit for  $142.32.6/24$

BGP table:

Prefix	Exit/Gateway
--------	--------------

✓  $142.32.6/24$   $R_1$   
true for all

IGP table:

IP addr	Next Hop
$\vdots$	$\vdots$

(say  $R_9$ )


$142.32.6/24$   
↑

→ Recursive Lookup.  
On receiving a packet, it finds out dest. prefix; looks for gateway in BGP table ( $R_1$ ) & then looks for next hop in the IGP table ( $R_{11}$ ).  $R_{11}$  does a similar thing and so on.

## (iii) Tagged IGP:

Internal routers may not be BGP speakers here. But IGP allows add<sup>n</sup> of "tags" in the ads.

$R_1$  inserts into its own IGP table:


 142.32.6/24 prefix tag  $R_1$   $\rightarrow$  gateway router.

This info is propagated to all routers in AS using IGP

Now @  $R_9$ :

IGP Table:

	Dest $R_1$ $R_2$ ...	Next :	Tag : $R_1$ $R_2$	Cost 5 10 :
announced by $R_1$ $\leftarrow$	142.32.6/24			
announced by $R_2$ $\leftarrow$	142.32.6/24			

Thus,  $R_9$  looks up for the prefix & finds closest match. (cost-wise)  $\nearrow$  hot-potato routing.